Sonoma Technology, Inc.

1360 Redwood Way, Suite C
Petaluma, CA 94954-1169
707/665-9900
FAX 707/665-9800
www.sonomatech.com

CALIFORNIA REGIONAL $PM_{10}/PM_{2.5}$ AIR QUALITY STUDY (CRPAQS) MANAGEMENT OF ANCHOR SITE DATA

FINAL REPORT
STI-999242-2087-FR

By:
Hilary R. Hafner
Nicole P. Hyslop
Cynthia N. Green
Sonoma Technology, Inc.
1360 Redwood Way, Suite C
Petaluma, CA 94954-1169

Prepared for:
The San Joaquin Valleywide Air Pollution Study Agency
c/o Karen Magliano
California Air Resources Board
1001 I Street
Sacramento, CA 95812

May 30, 2003

This page is intentionally blank.

ACKNOWLEDGMENTS

We thank the The San Joaquin Valleywide Air Pollution Study Agency for sponsoring this project and Karen Magliano of the California Air Resources Board, CRPAQS Project Manager, for her oversight. We acknowledge the considerable contributions of the following individuals:

- Measurement Experts–
  - Dr. Judith Chow (Desert Research Institute)
  - Dr. L. Willard Richards (Sonoma Technology, Inc.)
  - Dr. Susanne Hering (Aerosol Dynamics, Inc.)
  - Dr. Dennis Fitz (CE-CERT)
  - Mark Stoelting (STI)
- Dr. Ann E. (Beth) Wittig, STI's Field Manager
- Field Technicians – Matt Meyers, Wanda Sullivan, and Christopher Rumm (STI); Earle Wright (Harding ESE)
- Programming Assistance, STI – Eric Gray, Jeffrey Prouty, Steven Ludewig, and Dr. Patrick Ryan
- Database Management, Validation, and Documentation, STI – Cynthia Green, Carryl Hardy, and Erin Shields
- Data Validation and Documentation – Jerry Anderson, Courtney Gorin, Theresa O'Brien, Siana Alcorn, Will Richards, Liz Simon, and Duc Nguyen (STI); Mark Stoeltzenburg and Brent Kirby (ADI); Dennis Fitz (CE-CERT)
- Data Quality Summary Reports, STI – Steven Brown, Erin Shields, Courtney Gorin, and Jeffrey Prouty

We also acknowledge the helpful technical guidance of STI senior scientists, Dr. Donald Blumenthal and Dr. Paul Roberts.

This page is intentionally blank.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Section **Page**

TABLE OF CONTENTS (Concluded)

This page is intentionally blank.

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Concluded)

Figure **Page**

This page is intentionally blank.

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF TABLES (Concluded)

This page is intentionally blank.

LIST OF ABBREVIATIONS

ADI	Aerosol Dynamics, Inc.
agl	above ground level
API	application programming interface
ARB	California Air Resources Board
BAM	beta attenuation mass monitor
BC	black carbon
b_{sp}	particle scatter
CDF	comma-delimited field
CE-CERT	Center for Environmental Research and Technology
CO_2	carbon dioxide
CPC	condensation particle counter
CRPAQS	California Regional $PM_{10}/PM_{2.5}$ Air Quality Study
CV	coefficient of variation
DAP	data acquisition program
DAS	data acquisition system
DMA	differential mobility analyzer
DMC	data management center
DRI	Desert Research Institute
EC	elemental carbon
ESC	electrostatic classifier
FTP	file transfer protocol
GPT	gas-phase titration
HEPA	high efficiency particulate air
HNO_3	nitric acid
ID	identification code
ini	initialization
LPM	liters per minute
LQL	lower quantifiable limit
MOUDI	micro-orifice uniform deposit impactor
MSL	mean sea level
NDIR	non-dispersive infrared
NH_3	ammonia
NHO_3	nitric acid
NIR	near infrared
NIST	National Institute of Standards and Technology
NO	nitrogen oxide
NO_2	nitrogen dioxide
NO_x	oxides of nitrogen
NO_y	oxides of nitrogen
NPN	n-propylnitrate
O_3	ozone
OC	organic carbon
OCEC	organic carbon/elemental carbon continuous analyzer

LIST OF ABBREVIATIONS (Continued)

ODBC	Open DataBase Connectivity
OP code	operating code
OPC	optical particle counter
P	pressure
PAN	peroxyacetylnitrate
PI	principal investigator
PM_{10}	particulate matter less than 10 microns
$PM_{2.5}$	particulate matter less than 2.5 microns
PMS	Particle Measuring Systems
PSL	polystyrene latex
ppb	parts per billion
ppm	parts per million
ppt	parts per trillion
QA	quality assurance
QC	quality control
RH	relative humidity
RMS	root mean square
SJV	San Joaquin Valley
SLPM	standard liters per minute
SMPS	scanning mobility particle sizer
S/N	serial number
SO_2	sulfur dioxide
SOP	standard operating procedure
SQL	Structured Query Language
STI	Sonoma Technology, Inc.
SUVA	Freon 134a
T	temperature
TC	technical coordinator
TSI	Thermo-Systems, Inc.
UV	ultraviolet
WD	wind direction
WS	wind speed

Site Names and Codes

- ALT1 Altamont Pass
- ANGI Angiola
- ANGT Angiola Tower
  - ANG1 1-meter Angiola Tower
  - ANG50 50-meter Angiola Tower
  - ANG95 95-meter Angiola Tower
- BAC Bakersfield
- BODB Bodega Bay

LIST OF ABBREVIATIONS (Concluded)

Site Names and Codes (Concluded)

BTI	Bethel Island
COP	Corcoran Patterson Avenue
EDW	Edwards Air Force Base
FST	Fresno Supersite
M14	Modesto 14th Street
SDP	Sacramento Del Paso
SJ4	San Jose 4th Street
SNFH	Sierra Nevada Foothills
WAG	Walnut Grove
WAGT	Walnut Grove Tower

This page is intentionally blank.

1. INTRODUCTION TO CRPAQS, ANCHOR SITE OPERATIONS, AND DATA MANAGEMENT

1.1 CRPAQS OBJECTIVES

The California Regional $PM_{10}/PM_{2.5}$ Air Quality Study (CRPAQS) is a multi-year program of meteorological and air quality monitoring, emission inventory development, data analysis, and air quality simulation modeling. The objectives of CRPAQS are to (1) improve the understanding of emissions and the dynamic atmospheric processes that influence particle formation and distribution; (2) develop and demonstrate methods useful to decision makers in formulating and comparing candidate control strategies for attaining the federal and state $PM_{10}/PM_{2.5}$ standards in central California; and (3) provide reliable means for estimating the impacts of control strategy options developed for $PM_{10}/PM_{2.5}$ on visibility, air toxics, and acidic aerosols and on attainment strategies for other regulated pollutants, notably ozone. Detailed field study objectives are found in the field program plan compiled by Desert Research Institute (DRI) (Watson et al., 1998).

Meeting these objectives requires an extensive, high-quality air quality and meteorological database. Air quality data for the study were obtained during a 14-month field program by a combination of full-scale anchor monitoring sites measuring both gaseous and aerosol species, supplemental satellite sites measuring aerosol species using portable monitors, and a backbone network of existing California Air Resources Board (ARB) and air pollution control district sites.

The field phase of CRPAQS included annual monitoring throughout the San Joaquin Valley (SJV) and surrounding regions from December 1, 1999, through February 3, 2001, and intensive measurements during fall and winter conditions when PM_{10} and $PM_{2.5}$ concentrations are highest. The intensive measurements were made during the fall and second winter of the annual program. This report describes the data management, which includes data collection, processing, validation, and delivery, of measurements made at the anchor sites and on two tall towers which were operated by Sonoma Technology, Inc. (STI).

1.2 OVERALL SUMMARY OF ANCHOR SITES

The anchor site network consisted of three core sites for the annual and eight-week winter intensive monitoring (located at Fresno, Bakersfield, and a non-urban location near Angiola in Tulare County), plus two additional annual and winter sites (in Sacramento and San Jose) with more limited instrumentation. Six satellite sites were upgraded to anchor sites during the winter period, and one satellite site was upgraded to an anchor site during the summer period. **Figure 1-1** shows a map of the site locations. More detailed site information is provided in **Table 1-1.**

Arrays of continuous particle and gas monitors were deployed at anchor sites, with averaging times of 1 to 60 minutes. A complete list of measurements made during the annual, summer, fall, and winter portions of the study, periods of operation of each instrument by site, specific measurement objectives of the anchor sites, organization charts for the anchor sites, responsibilities, contact information, and other site information is discussed in detail in the STI field report (Wittig et al., 2003).

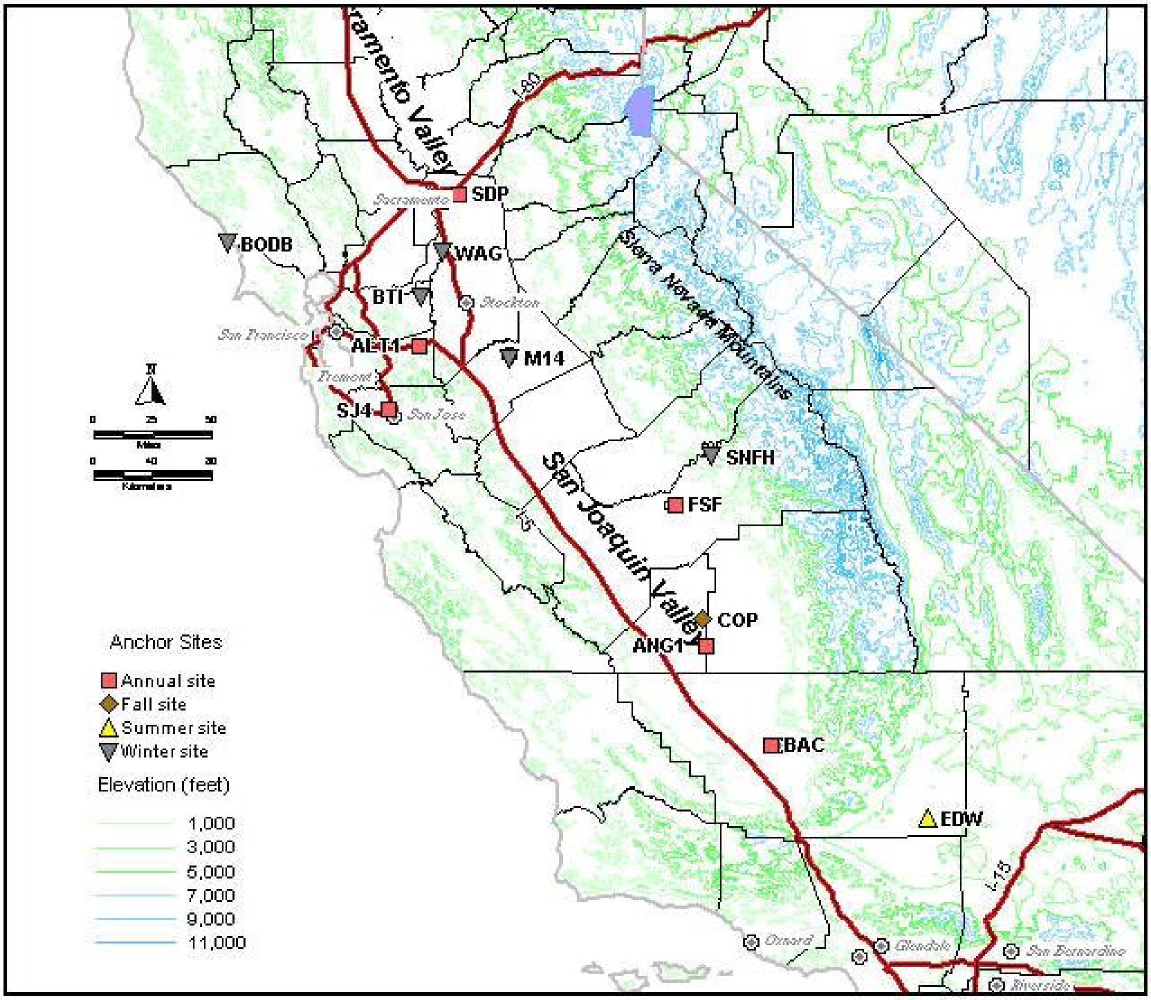


Figure 1-1. CRPAQS anchor sites. Site information and definitions of abbreviations are provided in Table 1-1.

1.3 SONOMA TECHNOLOGY, INC. DATA MANAGEMENT ROLE

STI's data management activities were designed to provide an accurate, precise, and complete data set for the anchor site measurements. In addition, the data management activities include documentation of all data collection, processing, validation, and archiving steps so that the data set can be evaluated and modified, if needed, at a later time. Ms. Hilary Hafner served as the Data Manager. Dr. Paul Roberts, the Technical Coordinator (TC), advised Ms. Hafner on data management issues and oversaw data management activities. Ms. Cynthia Green, and later Ms. Carryl Hardy and Ms. Erin Shields, served as the key staff members responsible for daily data management activities. Ms. Nicole Hyslop managed and advised staff regarding data validation issues.

Table 1-1. CRPAQS anchor sites.

ID	Site	Address	Latitude [a] (degrees)	Longitude [a] (degrees)	MSL Elevation[b] (meters)	Study Period[d] A	S	F	W
ALT1	Altamont Pass	Flynn Road Exit, I-580	37.718	-121.660	350[c]	x			
ANGI	Angiola trailer	36078 4th Avenue, Corcoran	35.948	-119.538	60	x			x
ANGT	Angiola tower (3 heights: ANG1, ANG50, ANG95)	36078 4th Avenue, Corcoran	35.948	-119.538	61, 110, 155	x			x
BAC	Bakersfield, California Avenue	5558 California Avenue #430, Bakersfield	35.357	-119.063	119	x			x
BODB	Bodega Bay	Bodega Marine Lab, 2099 Westside Road, Bodega Bay	38.319	-123.073	17				x
BTI	Bethel Island	5551 Bethel Island Road, Bethel Island	38.006	-121.642	2				x
COP	Corcoran, Patterson	1520 Patterson Ave., Corcoran	36.102	-119.566	63			x	
EDW	Edwards Air Force Base	Rawinsonde Road, Edwards Air Force Base	34.929	-117.904	724		x		
FSF	Fresno Supersite	3425 First Street, Fresno	36.781	-119.772	97	x			
M14	Modesto, 14th St.	814 14th Street, Modesto	37.634	-120.994	28				x
SDP	Sacramento, Del Paso Manor	2700 Maryal Drive, Sacramento	36.614	-121.368	26	x			x
SJ4	San Jose, 4th Street	120 N. 4th Street, San Jose	37.340	-121.889	26	x			
SNFH	Sierra Nevada Foothills	31955 Auberry Road, Auberry	37.063	-119.496	589[c]				x
WAG/ WAGT	Walnut Grove tower (10 m msl, 245 m msl)	KCRA-TV tower, Walnut Grove	38.264	-121.491	10, 245				x

[a] Coordinates are referenced to the NAD83 map datum .
[b] MSL elevation for all sites besides mountain sites is +/- 1 meter.
[c] MSL elevation for mountain sites is +/- 5 meters.
[d] A = Annual, S = Summer, F = Fall, W = Winter. Annual sites were operated for the entire study. Marks under S, F, or W for sites operated during the annual measurement period indicate that additional instruments were added to the sites for the seasonal studies.

Data from continuous monitors flowed from onsite instruments to onsite data acquisition systems (DAS) to the STI data management center (DMC) computer. Data from individual instruments included time-averaged values plus associated quality-control (QC) data. At the DMC, the data were loaded into a relational database, plotted, reviewed for consistency and operational status each day, and eventually validated and delivered to the CRPAQS Data Manager at ARB. The daily data plots were also uploaded to a restricted web site for access by the TC, measurements experts, and field staff. The purpose of the daily review at the DMC was to ensure proper operation of the field equipment and to provide feedback to the field operators when potential problems were identified. In addition, the field operators or field manager performed daily reviews of the data. An electronic log was kept of all actions taken on the individual values in the data set.

STI's responsibility for meeting the overall CRPAQS objectives for the anchor site measurements included testing, installing, and operating instruments in the field and collecting, managing, validating, and delivering data from these instruments. As a part of this endeavor, STI assigned a full-time field manager, several field technicians, and a data management technician. The PI, TC, data manager, programmer, and many other technical professionals at STI were also involved daily in the process. A unique aspect of the study was the use of measurement experts who provided guidance on the operation of the instruments maintained at each of the anchor sites as well as data validation direction and review. **Table 1-2** lists the measurement experts.

Table 1-2. Measurement experts.

Measurement/Instrument	Expert
Aethalometers	Dr. Beth Wittig (STI)
Continuous carbon (OCEC)	Dr. Paul Roberts (STI)
Continuous PM_{10} and $PM_{2.5}$ mass, VOC, and ozone	Dr. Paul Roberts (STI)
Continuous sulfate, nitrate, and particle-size	Dr. Susanne Hering (ADI)
Data acquisition systems	Mark Stoelting (STI)
Filter samplers and MOUDIs	Dr. Judith Chow (DRI)
Nephelometers	Dr. L. Willard Richards (STI)
PAN/NO_2, NO_y, and nitric acid	Mr. Dennis Fitz (CE-CERT)
SO_2	Earle Wright (Harding ESE) [Mr. Wright is currently affiliated with TMSI.]

Table 1-3 lists the parameters that STI validated and delivered. The table includes the averaging periods delivered.

1.4 SUMMARY OF OTHER REPORTS

There are several different types of information that are available to further interpret data collected during the CRPAQS field study. During the field study several formal documents were submitted to the funding agency. These documents contain detailed information about the field measurements at all phases of the study, including planning, operations, and reporting.

Planning

- Aerometric Monitoring Program Plan for the California Regional $PM_{2.5}$/PM_{10} Air Quality Study (DRI 9801.1D5) (Watson et al., 1998)
- A Proposal for the California Regional $PM_{2.5}$/PM_{10} Air Quality Study (STI 798900)
- California Regional PM_{10}/$PM_{2.5}$ Air Quality Study: Objectives and Associated Data Analysis and Modeling Approaches (Magliano et al., 1999)

Table 1-3. Parameters measured, validated, and delivered to ARB by STI. Averaging periods (Interval) are also included. Abbreviations are defined in the List of Abbreviations at the beginning of this report.

Instrument	Measurement	Interval[a]	Sites[b]																
			ALT1	ANGI	ANG95	ANG50	ANG1	BAC	BTI	BODB	COP	EDW	FSF	M14	SDP	SJ4	SNFH	WAG	WGT
Aethalometer	BC(1,7)[c]	5, 60		x	x[d]			x	x	x[d]	x[e]	x	x[d]	x[e]	x	x	x	x[d]	x[d]
BAM	PM_{10} mass	60		x				x	x		x	x					x		
BAM	$PM_{2.5}$ mass	60	x	x				x	x		x	x			x	x	x		
Meteorological	WS, WD, T, RH																x		
Nitric Acid	Nitric acid	5, 60		x													x		
Nephelometer	b_{sp}, RH	5, 60		x	x	x	x	x	x		x	x		x	x	x	x	x	x
Nitrate	Nitrate	10, 60		x	x			x	x		x					x	x	x	x
NO_y	NO, NO_y	5, 60		x	x			x	x				x				x		
Ozone	Ozone	5, 60		x	x			x	x								x		
OCEC	OC, OCEC	60		x				x	x										
PAN/NO_2	PAN, NO_2	5, 60		x				x	x								x		
Climet OPC	particle counts	5, 60		x	x	x			x										
PMS Lasair OPC	particle counts	5, 60		x					x										
SMPS	particle counts	5, 60		x															
SO_2	SO_2	5, 60						x											
Sulfate	Sulfate	10, 60			x			x											

[a] Interval = sample averaging interval delivered to ARB CRPAQS database in minutes (i.e., 5, 10, and 60 minutes)
[b] Site codes are defined in Table 1-1.
[c] BC(1,7) = black carbon (1 and 7 wavelengths)
[d] Only BC(1)
[e] Only BC(7)

Operations

- Quality Integrated Work Plan for the California Regional $PM_{10}/PM_{2.5}$ Air Quality Study Continuous and Filter Air Quality Measurements (Wittig et al., 2000)
- Health and Safety Plan for the California Regional $PM_{10}/PM_{2.5}$ Air Quality Study (Wittig et al., 1999)
- Audit Reports – These reports were compiled after each audit by the Quality Assurance (QA) Manager. They contain the numerical results of each audit and summaries of the system findings and performance results.
- Standard Operating Procedures (SOPs) – These SOPs were prepared for each instrument and are contained in the STI anchor site operations report (Wittig et al., 2003).
- ENSR's CRPAQS Site Documentation Report – This report contains descriptive site information, as well as photos, site diagrams, and maps (McDade, 2002).
- Field Management Report for Anchor Sites – This report describes the operations at the anchor sites, the site environments, and the instruments used and includes the SOPs (Wittig et al., 2003).

Data Quality

- Data Quality Summary Reports – These reports were compiled after data validation was completed to document data completeness, accuracy, precision, and lower quantifiable limits (Hyslop et al., 2003)

1.5 GUIDE TO THIS REPORT

This report documents STI's data management activities. Section 2 discusses data processing steps. Section 3 summarizes the review and application of calibration data to the raw data as well as the data validation steps. Section 4 describes STI's data archive. Section 5 provides a guide to the data quality summary reports. Section 6 lists references. To facilitate following the flow of data processing, calibration, and validation for a particular instrument, **Table 1-4** summarizes the location in this report of these steps for each instrument. The data quality summary reports are provided by (Hyslop et al., 2003).

Table 1-4. Guide to sections with data processing, calibration data review, and data validation discussions by instrument.

Manufacturer/Model	Instrument	Data Processing (Section 2)	Calibration Data Review (Section 3)	Data Validation (Section 3)
API Model 400A	Ozone	2.5.1	3.4.2	3.6.1
ThermoEnvironmental 42CY	NO_y	2.5.2	3.4.3	3.6.2
Radiance Research M903	Nephelometer	2.5.3	3.4.4	3.6.3
Met One 1020	BAM $PM_{2.5}$ & PM_{10}	2.5.4	3.4.5	3.6.4
Anderson Instruments AE-1X, AE-3X	Aethalometer	2.5.5	3.4.6	3.6.5
Rupprecht & Patashnick 5400	OCEC	2.5.6	3.4.7	3.6.6
ThermoEnvironmental 43S	SO_2	2.5.7	3.4.8	3.6.7
CE-CERT Gas Chromatograph	PAN/NO_2	2.5.8	3.4.9	3.6.8
ThermoEnvironmental 42CY	Nitric Acid	2.5.9	3.4.3	3.6.9
Rupprecht & Patashnick 5400N	Nitrate	2.5.10	3.4.10	3.6.10
Rupprecht & Patashnick 5400S	Sulfate	2.5.11	3.4.10	3.6.11
Climet Instruments CI-500 SPECTRO 0.3	OPC	2.5.13	3.4.11	3.6.12
PMS Lasair 1003	OPC	2.5.14	3.4.11	3.6.13
Thermo-Systems, Inc.	SMPS	2.5.12	3.4.12	3.6.14

This page is intentionally blank.

2. DATA PROCESSING PROCEDURES

The flow of data from the instruments to the STI CRPAQS database is discussed in detail in this section. Many acronyms and file names are used in this and later sections; these are summarized at the beginning of this report.

2.1 OVERVIEW

This field study challenged our data handling capabilities—more than 50 million records reside in the STI database consisting of 15 Gb of data and 5 Gb of log entries. A typical daily data file averaged 2 Mb. Each day, up to 90 time series plots were automatically generated and posted to a restricted web site. More than 375 air quality, meteorological, and operational parameters were collected and stored in the database. Time resolution for the various instruments included 1, 5, 10, and 60 minutes. Three instruments provided size-resolved aerosol counts (with 8, 16, and 53 size cuts), and one instrument provided measurements at 7 wavelengths.

Figure 2-1 illustrates overall daily operations during and after the field study. The data (from continuous monitors) flowed from the onsite instrument to the onsite DAS to the STI DMC computer. Data not yet stored at STI were automatically transferred electronically every night. Data from an individual instrument included time-averaged values plus associated QC information. A large hard drive at each site held more than two months of data to ensure that there would be no data loss in the event of polling problems. Once the DMC received the data each morning, the data were automatically loaded into a relational database, selected measurements were plotted, and the plots and data transfer were reviewed for consistency and operational status each day. Eventually the data were validated and delivered to the CRPAQS Data Manager at ARB. Daily operations were designed to ensure high-quality data.

2.2 DATA ACQUISITION

2.2.1 Daily Routine Data Acquisition

The CRPAQS DAS is a flexible, programmable data collection system designed by STI for air quality measurements. Data were collected on a standard, Pentium personal computer in ASCII files that could be read by any database parser, spreadsheet, or text editor. Both analog and serial data[1] were collected, averaged, and stored during 1-minute, 5-minute, 10-minute, or 60-minute data collection intervals, depending on the instrument. Digital and serial commands operated pumps and valves and conducted automated calibration operations.

[1] Both analog and serial signals from the instrument were collected because the digital (serial) signal was occasionally missing. When the digital signal was missing, the analog data were used to fill the time gap.

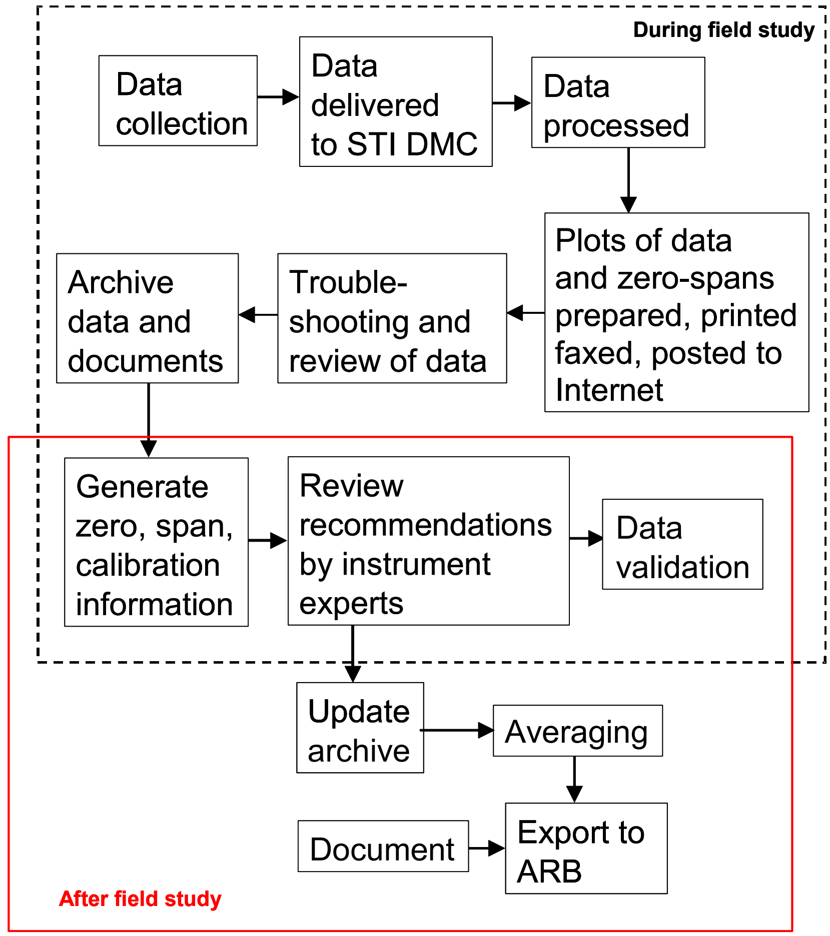


Figure 2-1. Overview of daily operations during and after the field study.

A single raw file was generated for each instrument each day and was stored in the on-site DAS. The DAS was able to parse any serial output directly into a 1-day and 3-day running data file (Recent.DAT). The parser was able to identify which particular fields of data to extract from the raw data file and copy into the running files. Typically, diagnostic data and instrument measurement values were parsed and copied into the running files. The benefit of this system was that the raw data files could be maintained for detailed information when needed, and the information maintained in the running file (and consequently in the STI CRPAQS archive) could be condensed and focused.

At midnight every night, the data acquisition program closed all data files for that day and truncated the Recent.DAT data file to a specified number of data records. The DAS then initiated a dial-up connection to the STI CRPAQS server. The Recent.DAT file was uploaded nightly, starting after midnight. The uploading times at different sites for these files were staggered to ensure files were transferred completely to an FTP file server at STI's DMC in Petaluma. The Recent.DAT file contained a total of three days of data: the most recent day and the two days previous. In the event that a nightly data file transfer was unsuccessful, the STI technician would either request the site operator to manually induce FTP transfer or dial into the DAS system using pcAnywhere (dial-up software) to initiate transmission of remote site files.

At STI, files were stored in a temporary directory to await processing. Starting at 0700 PST daily, the data files were automatically parsed and added to the database. Once the files were processed, they were stored in the “processed” folder. These daily Recent.DAT files were zipped and stored in the “processed_archive” directory at the end of each month.

A header record in the Recent.DAT file contained information about the DAS software version number and the data source (analog or serial data and channel number) for each instrument pulled from the raw files. The date/time stamp was the sampling end date and time for each measurement. Comma-delimited fields containing the raw data values for the various measurements followed the date/time stamp. There was also a field containing operating code (OP code) information for each parameter in the file (OP codes are discussed in Section 2.3.1). The last record on each line in the Recent.DAT file was reserved for a field containing comments manually entered by site operators (note: these comments were not tied to particular parameters, only to the date/time stamp). **Figure 2-2** shows an excerpt from an example Recent.DAT file.

(0602000714)_sj4 recent.dat - Notepad

File Edit Format Help

```
REM, CRPAQS Data Acquisition System v.00.03.02
REM, Instrument: Neph
REM, Source: Analog Channel 0
REM, Units: %
REM, Zero Offset: 0.000000
REM, Slope: 20.000000
REM, Source: Analog Channel 1
REM, Units: 1/Mm
REM, Zero Offset: 0.000000
REM, Slope: 200.000000
REM, Source: Analog Channel 2
REM, Units: 1/Mm
REM, Zero Offset: 0.000000
REM, Slope: 200.000000
REM, Source: Analog Channel 3
REM, Units: VDC
REM, Zero Offset: 0.000000
REM, Slope: 1.000000
REM, Instrument: Aeth
REM, Source: S4
REM, Database Items: Aeth_BC,Aeth_Flow,Aeth_sz,Aeth_sb,Aeth_rz,Aeth_rb,,,REM, Instrument: BAM25
REM, Source: S5
REM, Database Items: BAM25_MassConc,BAM25_Vol,,
HDR,SJ4,DATE/TIME,NephAnalog_RH,NephAnalog_RH_OP,NephAnalog_bsp,NephAnalog_bsp_OP,NephAnalog_Calib,NephAnalog_Calib_OP,NephAn
DAT,SJ4,05/30/2000 00:01:00,67.792969,0,41.341146,0,244.205729,0,3.849731,0,0.1580,0,6.60,0,0.02150,0,0.36150,0,0.02130,0,2.2
DAT,SJ4,05/30/2000 00:02:00,67.932129,0,41.691081,0,243.448893,0,3.842122,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:03:00,68.068034,0,42.317708,0,243.221029,0,3.840942,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:04:00,68.028157,0,41.349284,0,242.789713,0,3.841268,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:05:00,68.052572,0,41.227213,0,242.854818,0,3.845703,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:06:00,68.020833,0,42.130534,0,243.245443,0,3.840129,0,0.1640,0,6.50,0,0.02150,0,0.36070,0,0.02130,0,2.2
DAT,SJ4,05/30/2000 00:07:00,68.127441,0,42.814127,0,243.098958,0,3.840291,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:08:00,68.159180,0,5.818685,0,519.482422,0,4.257324,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:09:00,68.177897,0,48.177083,0,218.855794,0,3.832275,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:10:00,68.055013,0,43.416341,0,232.828776,0,3.839518,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:11:00,68.033040,0,42.960612,0,238.142903,0,3.844442,0,0.2160,0,6.60,0,0.02150,0,0.360,0,0.02130,0,2.217
DAT,SJ4,05/30/2000 00:12:00,68.079427,0,42.000325,0,240.763346,0,3.843994,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:13:00,68.182780,0,40.828450,0,242.285156,0,3.838297,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:14:00,68.158366,0,41.780599,0,243.025716,0,3.846151,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:15:00,68.138021,0,41.861979,0,243.709310,0,3.840128,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:16:00,68.179525,0,42.822265,0,243.823242,0,3.848185,0,0.2160,0,6.60,0,0.02150,0,0.35930,0,0.02130,0,2.2
DAT,SJ4,05/30/2000 00:17:00,68.186035,0,42.626953,0,244.335937,0,3.842855,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:18:00,68.204753,0,42.122396,0,244.091797,0,3.835205,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:19:00,68.291829,0,42.415364,0,243.538411,0,3.842895,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:20:00,68.315430,0,43.709310,0,243.945312,0,3.842611,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:21:00,68.257650,0,42.692057,0,243.937174,0,3.841187,0,0.150,0,6.60,0,0.02150,0,0.35870,0,0.02130,0,2.21
DAT,SJ4,05/30/2000 00:22:00,68.313802,0,43.180338,0,242.659505,0,3.841553,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:23:00,68.238118,0,42.643229,0,242.846680,0,3.840088,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:24:00,68.322754,0,44.73470,0,242.952474,0,3.836873,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:25:00,68.424479,0,43.872070,0,243.481445,0,3.841309,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:26:00,68.353678,0,42.919922,0,243.880208,0,3.851359,0,0.1590,0,6.50,0,0.02150,0,0.35810,0,0.02130,0,2.2
DAT,SJ4,05/30/2000 00:27:00,68.418783,0,42.513021,0,245.019531,0,3.840006,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:28:00,68.435059,0,43.196614,0,244.140625,0,3.846558,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:29:00,68.388672,0,44.042969,0,244.547526,0,3.836507,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
DAT,SJ4,05/30/2000 00:30:00,68.427734,0,44.620768,0,244.759114,0,3.843750,0,M,9,M,9,M,9,M,9,M,9,M,9,M,9,M,9,
```

Figure 2-2. Excerpt from an example Recent.DAT file.

The DAS was also able to initiate automatic calibrations. These calibrations were scheduled in the initialization (INI) file and typically occurred after midnight. For those sites where automated instrument zero-span checks occurred on a nightly basis, a file called CalRecent.DAT was created every morning so that the most recent calibration could be monitored. This file contained data from 0000 to 0600 PST for the current day and was uploaded to the FTP server after the 0600 PST data record was written to the file. The file format was the same as that of Recent.DAT files, except that it contained only the most recent six hours of data.

Not all sites transferred the data nightly. The Altamont beta attenuation mass monitor (BAM) data (contained within a *.BAM file) was manually dumped into the FTP file server on a monthly basis where it was then automatically imported into the database at STI and processed. An excerpt from an example *.BAM file is shown in **Figure 2-3**.

J:\999242 CRPAQS DV Report\figures\example_data_files\(0727001111)_ALT 061600 BAM.bam

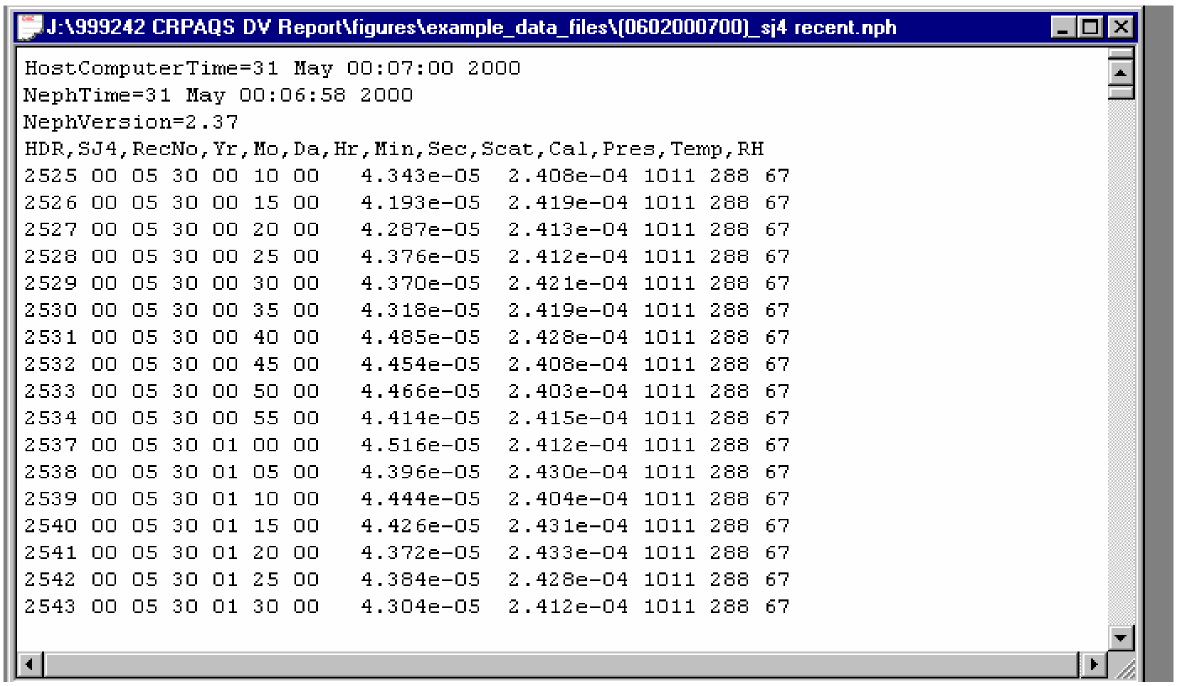
```
3

  Accessing Arrays


* 3

  Accessing Arrays


* 2

  Accessing Arrays



  Report for 06/03/2000 - Day 155    > BAM 1020 <    Station ID:  1

Channel                               01      02      03      04      05      06
Sensor               Conc    Qtot     XXX     XXX     XXX     XXX     XXX     XXX
Units                mg      m3       XXX     XXX     XXX     XXX     XXX     XXX
==========================================================================
05:00 ------------  0.016   0.842   00002   00002   00002   00002   00002   00002
06:00 ------------  0.010   0.839   00002   00002   00002   00002   00002   00002
07:00 ------------  0.019   0.843   00002   00002   00002   00002   00002   00002
08:00 ------------  0.011   0.839   00002   00002   00002   00002   00002   00002
09:00 ------------  0.012   0.829   00002   00002   00002   00002   00002   00002
10:00 ------------  0.008   0.823   00002   00002   00002   00002   00002   00002
11:00 ---------D--  0.007   0.825   00002   00002   00002   00002   00002   00002
12:00 ------------  0.012   0.818   00002   00002   00002   00002   00002   00002
13:00 ------------  0.012   0.822   00002   00002   00002   00002   00002   00002
14:00 ------------  0.007   0.817   00002   00002   00002   00002   00002   00002
```

Figure 2-3. Excerpt from *.BAM file.

DAS clocks were periodically checked against the National Institute of Standards and Technology (NIST) time from the NIST website (http://nist.time.gov/). Instrument clocks were then compared to DAS clocks and readjusted if necessary. The implication of time checks to data processing is that, in some cases, the DAS time stamp will not match the instrument time stamp when the raw data files from the instruments are compared to the Recent.DAT file. The DAS time stamp was used as the official time.

2.2.2 Data Acquisition via a "Backdoor"

For the nephelometer, a "backdoor" method of acquiring data was developed to obtain both the analog and digital data (and only one of these two data streams was transferable to the DAS). A different program, called GetNeph, polled the instrument daily after midnight, and created a one-day running raw data file, archived on the DAS hard drive, as well as a three-day Recent.NPH file (**Figure 2-4**). This Recent.NPH file was uploaded at the same time as the Recent.DAT file from the same site and was similarly processed at the STI DMC starting at 0700 PST. Some data (i.e., Sacramento) were downloaded by the site operators (T&B Systems) bi-weekly and sent to STI for processing.


J:\999242 CRPAQS DV Report\figures\example_data_files\(0602000700)_sj4 recent.nph

```
HostComputerTime=31 May 00:07:00 2000
NephTime=31 May 00:06:58 2000
NephVersion=2.37
HDR,SJ4,RecNo,Yr,Mo,Da,Hr,Min,Sec,Scat,Cal,Pres,Temp,RH
2525 00 05 30 00 10 00   4.343e-05  2.408e-04 1011 288 67
2526 00 05 30 00 15 00   4.193e-05  2.419e-04 1011 288 67
2527 00 05 30 00 20 00   4.287e-05  2.413e-04 1011 288 67
2528 00 05 30 00 25 00   4.376e-05  2.412e-04 1011 288 67
2529 00 05 30 00 30 00   4.370e-05  2.421e-04 1011 288 67
2530 00 05 30 00 35 00   4.318e-05  2.419e-04 1011 288 67
2531 00 05 30 00 40 00   4.485e-05  2.428e-04 1011 288 67
2532 00 05 30 00 45 00   4.454e-05  2.408e-04 1011 288 67
2533 00 05 30 00 50 00   4.466e-05  2.403e-04 1011 288 67
2534 00 05 30 00 55 00   4.414e-05  2.415e-04 1011 288 67
2537 00 05 30 01 00 00   4.516e-05  2.412e-04 1011 288 67
2538 00 05 30 01 05 00   4.396e-05  2.430e-04 1011 288 67
2539 00 05 30 01 10 00   4.444e-05  2.404e-04 1011 288 67
2540 00 05 30 01 15 00   4.426e-05  2.431e-04 1011 288 67
2541 00 05 30 01 20 00   4.372e-05  2.433e-04 1011 288 67
2542 00 05 30 01 25 00   4.384e-05  2.428e-04 1011 288 67
2543 00 05 30 01 30 00   4.304e-05  2.412e-04 1011 288 67
```

Figure 2-4. Excerpt from *.NPH file.

2.3 OVERALL PARAMETER PROCESSING DETAILS

2.3.1 Overview

The CRPAQS data management system uses Microsoft Access 97 as a front end to access data stored in the Microsoft Structured Query Language (SQL) Server 2000. There are four data storage tables in the SQL server corresponding to the different averaging periods encountered during the study: 1-, 5-, 10-, and 60-minute tables. A table containing the notes compiled during the data validation process (Data Log) is also linked from the SQL server, as are other tables involved in describing the data records stored (e.g., parameters, sites, and equipment).

After the files were uploaded to the CRPAQS FTP server, the data were automatically imported into the STI database. Automated database parsers ingested all data within the file that had an assigned parameter identification code (parameter ID). The parameter ID was required for successful importation. The parameter IDs were included in the header section of the Recent.DAT file daily before midnight and every time the DAS was restarted or began recording data. Variables without an assigned database parameter name in the initialization (INI) file were automatically left out of the Recent.DAT file.

Three types of files were recognized and processed by the data system: *.DAT, *.NPH, and *.BAM. These files were directed to the appropriate parser, and the data were sorted and appended to a table corresponding to the averaging time (1-, 5-, 10-, or 60-minutes) of the parameter. The parameters collected and their averaging periods are summarized in Table 1-3. **Table 2-1** summarizes the data record in the STI database.

Table 2-1. STI database record format.

Item	Parameter Abbreviation in the Database	Description
Site identification	SiteID	Site number corresponding to values in the SiteTextID table of the database
Parameter identification	ParamID	Parameter number assigned by STI
Original time stamp	OrigTimeStamp	DAS time stamp, indicates the end time of measurement record
Modified time stamp	ModTimeStamp	Indicates start time of measurement record
Raw data value	RawValue	Unchanged, raw data value
Modified data value	DataValue	Originally set to RawValue; value may change following data validation
Operation code	OpCode	Instrument specific operation code; indicates status of instrument
Quality control code	QCCode	Data validation code (0-9, see Table 2-2)
Quality control level	QCLevel	Level of validation that has been applied
Averaging period (minutes)	AvgPeriod	Number of minutes in the averaging period

The fields directly imported from the Recent.DAT file were the original time stamp, raw data value, and OP code. All other fields were either calculated (modified time stamp and data value), translated (site ID, parameter ID), or initialized (QC code, QC level, averaging period).

An OP code that contained operational information about a specific parameter was assigned to that parameter and written to an adjacent field in the daily Recent.DAT file stored on the DAS. Following the importation of the Recent.DAT file into the database, these OP codes were parsed together with the corresponding parameter to the same record. OP codes indicated the presence of instrument warning flags or indicated that the recorded values were part of a calibration routine. For some instruments, OP codes were automatically translated to corresponding QC codes as part of the import process. A table of OP codes and their QC code translation is shown in **Table 2-2**. During data validation (Section 3), technicians verified all QC codes.

Table 2-2. OP code-to-QC code translation

Equipment ID	OpCode	OP Code Description[a]	QC Code	QC Description
All equipment	0	Valid	0	Valid
NO_y	1	NO Span	2	Calibration/Instrument check
NO_y	2	GPT (NO_2/NO) converter check	2	Calibration/Instrument check
NO_y	3	NPN converter check	2	Calibration/Instrument check
NO_y	4	NH_3 converter check	2	Calibration/Instrument check
Ozone	4	O_3 span check	2	Calibration/Instrument check
NO_y	5	Zero air check	2	Calibration/Instrument check
Ozone	5	Zero air check	2	Calibration/Instrument check
NO_y	6	Matrix air check	2	Calibration/Instrument check
BAM PM	7	Instrument error	0	Valid
NO_y	7	Instrument error	3	Equipment problem
Ozone	7	Instrument error	3	Equipment problem
OCEC	7	Instrument error	3	Equipment problem
NO_y	8	HNO_3 gas	2	Calibration/Instrument check
All equipment	9	Missing	9	Missing

[a] Abbreviations are defined at the beginning of this report.

The modified time stamp field of the data record was created during the import process, using time stamp adjustment information shown in **Table 2-3**. The time adjustment value is subtracted from the original time stamp value of each record to give the actual start time of the sample record. For example, if the OCEC instrument began sampling at 0300 PST, the instrument collected air for one hour and then analyzed the sample for an additional hour. The instrument then reported a value at 0500 PST. Thus, 120 minutes are subtracted from the reported time to bring the data to begin time.

The modified data value is initially set to the raw value but with a defined precision. Precisions were selected based on the precision the instrument was capable of producing. Unlike the data value, which may change following data validation, the raw value is never changed and has no set precision, remaining the same as it reads in the Recent.DAT file.

Table 2-3. Time stamp modifications to convert DAS-stored data to actual start time of the measurement.

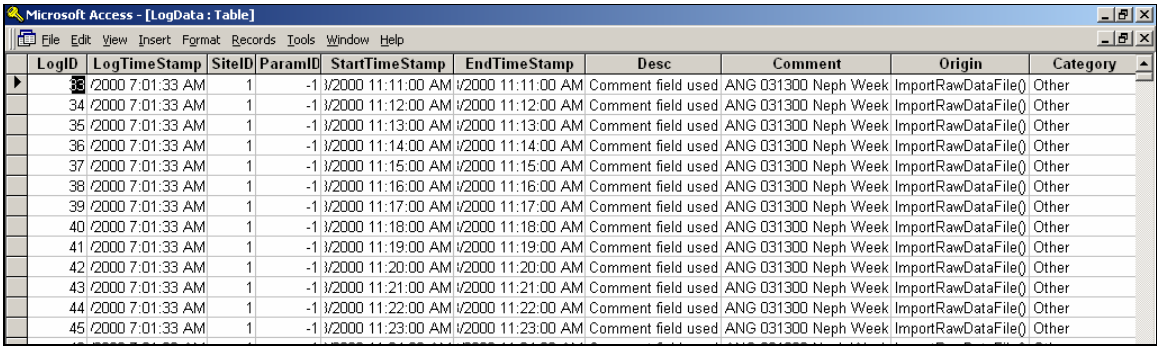
Instrument	Averaging Period (minutes)	Time Adjustment (minutes)
Aethalometer	5	-5
BAM PM	60	-60
Calibrator	1	-1
Climet OPC	5	-5
Meteorology (5 minute)	5	-5
Meteorology (60 minute)	60	-60
Nitric Acid	1	-1
Nephelometer (analog)	1	-1
Nephelometer (digital)	5	-5
Nitrate	10	-10
NO_y	1	-1
Ozone	1	-1
OCEC	60	-120
PAN/NO_2	1	-1
PMS OPC	5	-5
SMPS	5	-5
SO_2	1	-1
Sulfate	10	-10

2.3.2 Log Files

Three types of log files were generated following importation of the data files into the data management system: LogData files, LogEvent files, and LogImport files.

LogData

The LogData table stored the comment field that accompanied the daily Recent.DAT file. Site operators entered these comments (e.g., during manual calibrations) that described issues pertaining to the data during the indicated time periods. Comments were most often used to document periods when an instrument was undergoing calibration or maintenance procedures. Comments were not tied to a unique parameter in the database and were only reported with a site ID and start and end time stamps. **Figure 2-5 and Table 2-4** illustrate the structure of the LogData table. The LogData information was valuable as a reference in daily QC checks of the data and during data validation, and relevant information in the table was compiled during data validation procedures.



Microsoft Access - [LogData : Table]

File Edit View Insert Format Records Tools Window Help

LogID	LogTimeStamp	SiteID	ParamID	StartTimeStamp	EndTimeStamp	Desc	Comment	Origin	Category
33	/2000 7:01:33 AM	1	-1	/2000 11:11:00 AM	/2000 11:11:00 AM	Comment field used	ANG 031300 Neph Week	ImportRawDataFile()	Other
34	/2000 7:01:33 AM	1	-1	/2000 11:12:00 AM	/2000 11:12:00 AM	Comment field used	ANG 031300 Neph Week	ImportRawDataFile()	Other
35	/2000 7:01:33 AM	1	-1	/2000 11:13:00 AM	/2000 11:13:00 AM	Comment field used	ANG 031300 Neph Week	ImportRawDataFile()	Other
36	/2000 7:01:33 AM	1	-1	/2000 11:14:00 AM	/2000 11:14:00 AM	Comment field used	ANG 031300 Neph Week	ImportRawDataFile()	Other
37	/2000 7:01:33 AM	1	-1	/2000 11:15:00 AM	/2000 11:15:00 AM	Comment field used	ANG 031300 Neph Week	ImportRawDataFile()	Other
38	/2000 7:01:33 AM	1	-1	/2000 11:16:00 AM	/2000 11:16:00 AM	Comment field used	ANG 031300 Neph Week	ImportRawDataFile()	Other
39	/2000 7:01:33 AM	1	-1	/2000 11:17:00 AM	/2000 11:17:00 AM	Comment field used	ANG 031300 Neph Week	ImportRawDataFile()	Other
40	/2000 7:01:33 AM	1	-1	/2000 11:18:00 AM	/2000 11:18:00 AM	Comment field used	ANG 031300 Neph Week	ImportRawDataFile()	Other
41	/2000 7:01:33 AM	1	-1	/2000 11:19:00 AM	/2000 11:19:00 AM	Comment field used	ANG 031300 Neph Week	ImportRawDataFile()	Other
42	/2000 7:01:33 AM	1	-1	/2000 11:20:00 AM	/2000 11:20:00 AM	Comment field used	ANG 031300 Neph Week	ImportRawDataFile()	Other
43	/2000 7:01:33 AM	1	-1	/2000 11:21:00 AM	/2000 11:21:00 AM	Comment field used	ANG 031300 Neph Week	ImportRawDataFile()	Other
44	/2000 7:01:33 AM	1	-1	/2000 11:22:00 AM	/2000 11:22:00 AM	Comment field used	ANG 031300 Neph Week	ImportRawDataFile()	Other
45	/2000 7:01:33 AM	1	-1	/2000 11:23:00 AM	/2000 11:23:00 AM	Comment field used	ANG 031300 Neph Week	ImportRawDataFile()	Other

Figure 2-5. Excerpt from LogData table contained in the database.

Table 2-4. Summary of LogData table structure.

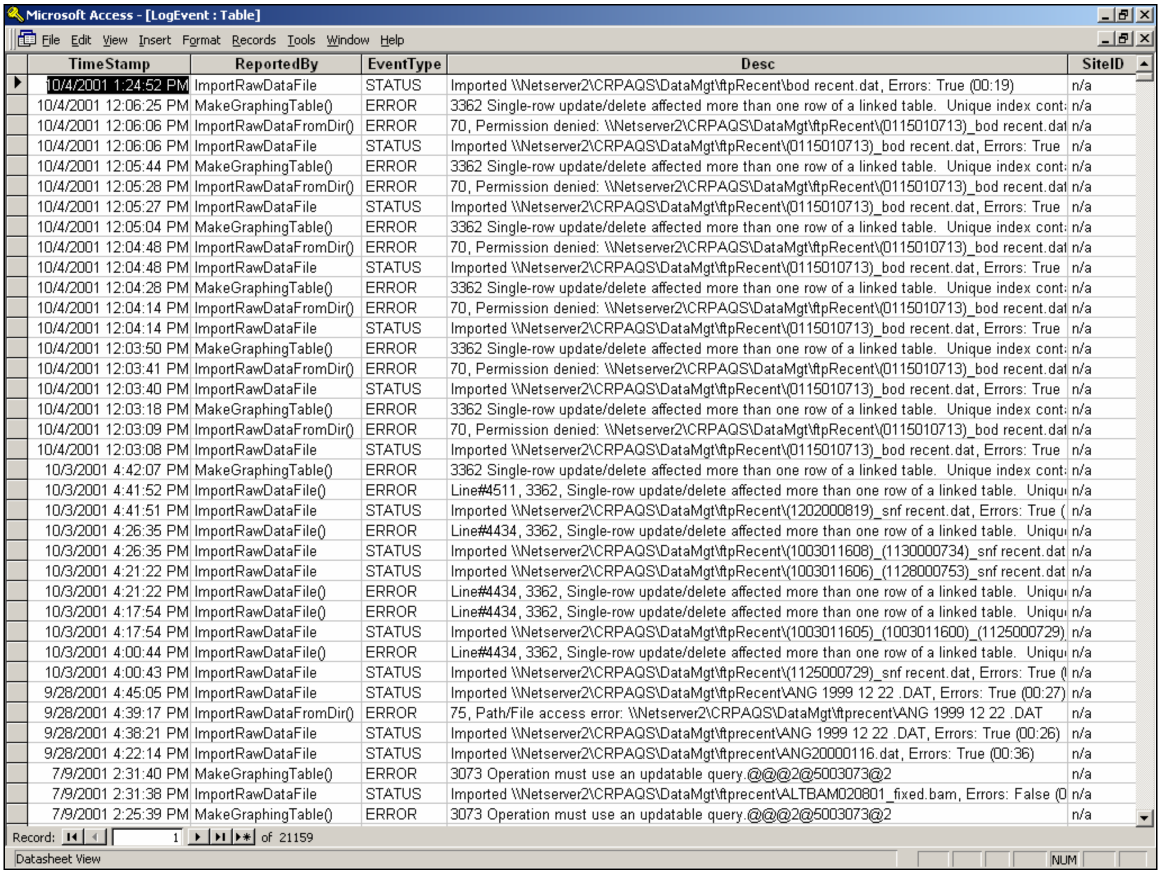
Abbreviation in the Database	Description
LogID	Auto numbered and unique ID
LogTimeStamp	Time when log file was generated by system
SiteID	No sites=-1; all sites=0; x= numeric SiteID
ParamID	"-1", corresponded to the parameter ID of the comment field
StartTimeStamp	Start time for comment
EndTimeStamp	End time for comment
Desc[a]	"Comment field used during data acquisition"
Comment	Description of what happened (and why) to the data value
Origin[a]	"ImportRawDataFile"
Category[a]	"Other"

[a] The Desc, Origin, and Category fields contained the default comments described because other components of the system were not implemented in this version of the database.

LogEvent

The LogEvent table **(Figure 2-6 and Table 2-5**) contained information describing the status of some data management processes, such as the auto-import of the raw data files, SurfDat (data validation tool described in Section 3) import, creation of graphing tables, and creation of *.GIFs for the web. The LogEvent information was referenced most often during troubleshooting procedures to ensure data were imported into the database and to confirm that certain processes had run to completion. When processes failed (e.g., ODBC[2] call failures), an error log was generated.

[2] Open Data Base Connectivity (ODBC) is an Application Programming Interface (API) that allows a programmer to abstract a program from a database. When writing code to interact with a database, the programmer adds code that talks to a particular database using a proprietary language (e.g., MS Access, Fox, Oracle). When programming to interact with ODBC, the programmer only needs to use the ODBC language (a combination of ODBC API function calls and the SQL language).



Microsoft Access - [LogEvent : Table]

File Edit View Insert Format Records Tools Window Help

TimeStamp	ReportedBy	EventType	Desc	SiteID
10/4/2001 1:24:52 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat, Errors: True (00:19)	n/a
10/4/2001 12:06:25 PM	MakeGraphingTable()	ERROR	3362 Single-row update/delete affected more than one row of a linked table. Unique index cont	n/a
10/4/2001 12:06:06 PM	ImportRawDataFromDir()	ERROR	70, Permission denied: \\Netserver2\CRPAQS\DataMgt\ftpRecent\(0115010713)_bod recent.dat	n/a
10/4/2001 12:06:06 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftpRecent\(0115010713)_bod recent.dat, Errors: True	n/a
10/4/2001 12:05:44 PM	MakeGraphingTable()	ERROR	3362 Single-row update/delete affected more than one row of a linked table. Unique index cont	n/a
10/4/2001 12:05:28 PM	ImportRawDataFromDir()	ERROR	70, Permission denied: \\Netserver2\CRPAQS\DataMgt\ftpRecent\(0115010713)_bod recent.dat	n/a
10/4/2001 12:05:27 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftpRecent\(0115010713)_bod recent.dat, Errors: True	n/a
10/4/2001 12:05:04 PM	MakeGraphingTable()	ERROR	3362 Single-row update/delete affected more than one row of a linked table. Unique index cont	n/a
10/4/2001 12:04:48 PM	ImportRawDataFromDir()	ERROR	70, Permission denied: \\Netserver2\CRPAQS\DataMgt\ftpRecent\(0115010713)_bod recent.dat	n/a
10/4/2001 12:04:48 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftpRecent\(0115010713)_bod recent.dat, Errors: True	n/a
10/4/2001 12:04:28 PM	MakeGraphingTable()	ERROR	3362 Single-row update/delete affected more than one row of a linked table. Unique index cont	n/a
10/4/2001 12:04:14 PM	ImportRawDataFromDir()	ERROR	70, Permission denied: \\Netserver2\CRPAQS\DataMgt\ftpRecent\(0115010713)_bod recent.dat	n/a
10/4/2001 12:04:14 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftpRecent\(0115010713)_bod recent.dat, Errors: True	n/a
10/4/2001 12:03:50 PM	MakeGraphingTable()	ERROR	3362 Single-row update/delete affected more than one row of a linked table. Unique index cont	n/a
10/4/2001 12:03:41 PM	ImportRawDataFromDir()	ERROR	70, Permission denied: \\Netserver2\CRPAQS\DataMgt\ftpRecent\(0115010713)_bod recent.dat	n/a
10/4/2001 12:03:40 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftpRecent\(0115010713)_bod recent.dat, Errors: True	n/a
10/4/2001 12:03:18 PM	MakeGraphingTable()	ERROR	3362 Single-row update/delete affected more than one row of a linked table. Unique index cont	n/a
10/4/2001 12:03:09 PM	ImportRawDataFromDir()	ERROR	70, Permission denied: \\Netserver2\CRPAQS\DataMgt\ftpRecent\(0115010713)_bod recent.dat	n/a
10/4/2001 12:03:08 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftpRecent\(0115010713)_bod recent.dat, Errors: True	n/a
10/3/2001 4:42:07 PM	MakeGraphingTable()	ERROR	3362 Single-row update/delete affected more than one row of a linked table. Unique index cont	n/a
10/3/2001 4:41:52 PM	ImportRawDataFile()	ERROR	Line#4511, 3362, Single-row update/delete affected more than one row of a linked table. Uniqu	n/a
10/3/2001 4:41:51 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftpRecent\(1202000819)_snf recent.dat, Errors: True (	n/a
10/3/2001 4:26:35 PM	ImportRawDataFile()	ERROR	Line#4434, 3362, Single-row update/delete affected more than one row of a linked table. Uniqu	n/a
10/3/2001 4:26:35 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftpRecent\(1003011608)_(1130000734)_snf recent.dat	n/a
10/3/2001 4:21:22 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftpRecent\(1003011606)_(1128000753)_snf recent.dat	n/a
10/3/2001 4:21:22 PM	ImportRawDataFile()	ERROR	Line#4434, 3362, Single-row update/delete affected more than one row of a linked table. Uniqu	n/a
10/3/2001 4:17:54 PM	ImportRawDataFile()	ERROR	Line#4434, 3362, Single-row update/delete affected more than one row of a linked table. Uniqu	n/a
10/3/2001 4:17:54 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftpRecent\(1003011605)_(1003011600)_(1125000729)	n/a
10/3/2001 4:00:44 PM	ImportRawDataFile()	ERROR	Line#4434, 3362, Single-row update/delete affected more than one row of a linked table. Uniqu	n/a
10/3/2001 4:00:43 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftpRecent\(1125000729)_snf recent.dat, Errors: True (	n/a
9/28/2001 4:45:05 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftpRecent\ANG 1999 12 22 .DAT, Errors: True (00:27)	n/a
9/28/2001 4:39:17 PM	ImportRawDataFromDir()	ERROR	75, Path/File access error: \\Netserver2\CRPAQS\DataMgt\ftprecent\ANG 1999 12 22 .DAT	n/a
9/28/2001 4:38:21 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftprecent\ANG 1999 12 22 .DAT, Errors: True (00:26)	n/a
9/28/2001 4:22:14 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftprecent\ANG20000116.dat, Errors: True (00:36)	n/a
7/9/2001 2:31:40 PM	MakeGraphingTable()	ERROR	3073 Operation must use an updatable query.@@@2@5003073@2	n/a
7/9/2001 2:31:38 PM	ImportRawDataFile	STATUS	Imported \\Netserver2\CRPAQS\DataMgt\ftprecent\ALTBAM020801_fixed.bam, Errors: False (0	n/a
7/9/2001 2:25:39 PM	MakeGraphingTable()	ERROR	3073 Operation must use an updatable query.@@@2@5003073@2	n/a

Record: 1 of 21159

Datasheet View NUM

Figure 2-6. Excerpt from LogEvent table contained in the database.

Table 2-5. Summary of LogEvent table structure.

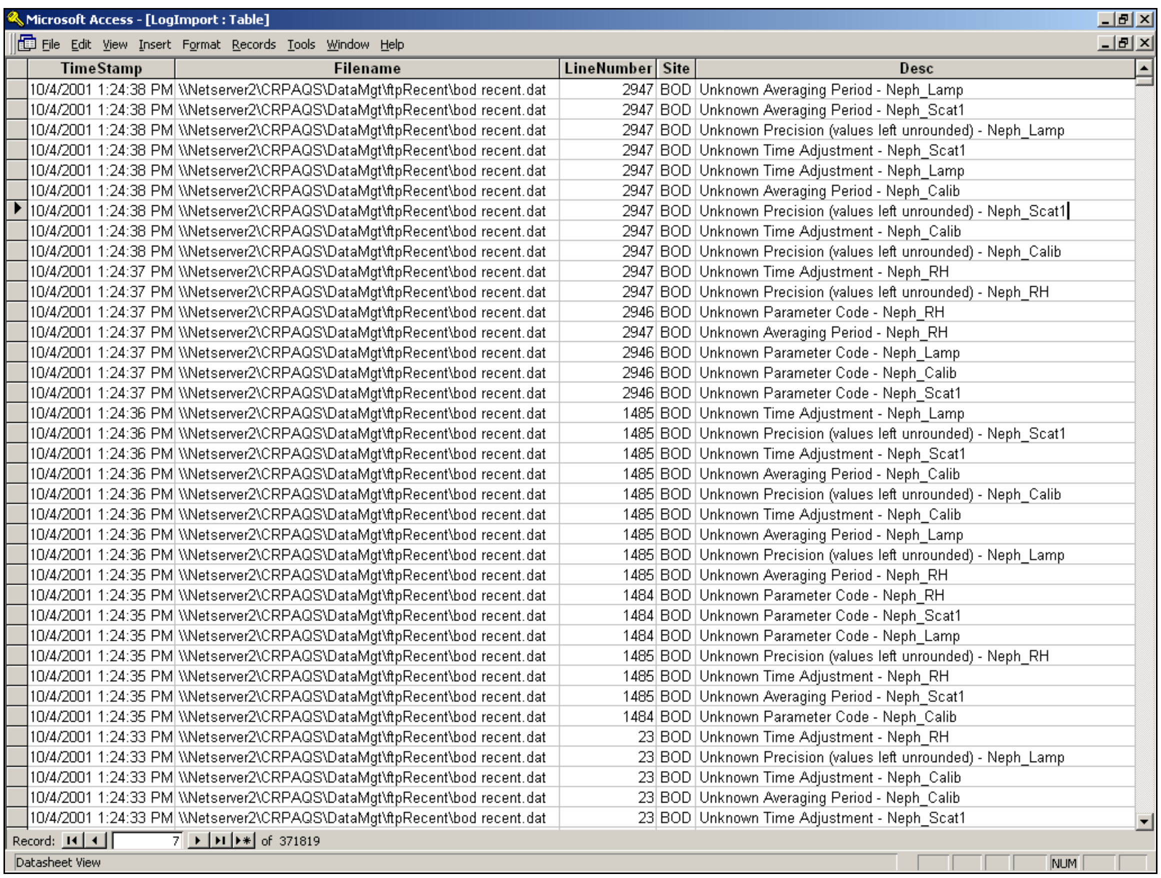
Abbreviation in the Database	Description
TimeStamp	Time when event occurred
Reported By	Where the event or database process took place; e.g.,“ImportRawDataFile” or “MakeGraphingTable”
EventType	“ERROR” or “STATUS”
Desc	Describes the type of error or status of the event
Site ID	Indicates for which site error or status occurred (usually “n/a” but an appropriate SiteID was provided when there were errors in the routine to make GIF files for posting to the website)

LogImport

The LogImport table (see **Table 2-6 and Figure 2-7**) was used to indicate line-specific problems with the incoming data files during the import process. This table was referenced when import problems were encountered and would often pinpoint errors in the format of the incoming data files.

Table 2-6. Summary of LogImport table structure.

Abbreviation in the Database	Description
TimeStamp	Time of import
Line number	Line number in file where problem occurs
Desc	Description of import problem (e.g. unknown parameter, averaging period, precision, etc.; usually due to the parameter not being specified correctly, or at all, in the SiteEquipParam table)
Filename	Name of raw data (.DAT) file and location
Site	Site corresponding to imported data file (three-letter code, SiteTextID)


Microsoft Access - [LogImport : Table]
File Edit View Insert Format Records Tools Window Help

TimeStamp	Filename	LineNumber	Site	Desc
10/4/2001 1:24:38 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2947	BOD	Unknown Averaging Period - Neph_Lamp
10/4/2001 1:24:38 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2947	BOD	Unknown Averaging Period - Neph_Scat1
10/4/2001 1:24:38 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2947	BOD	Unknown Precision (values left unrounded) - Neph_Lamp
10/4/2001 1:24:38 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2947	BOD	Unknown Time Adjustment - Neph_Scat1
10/4/2001 1:24:38 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2947	BOD	Unknown Time Adjustment - Neph_Lamp
10/4/2001 1:24:38 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2947	BOD	Unknown Averaging Period - Neph_Calib
10/4/2001 1:24:38 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2947	BOD	Unknown Precision (values left unrounded) - Neph_Scat1
10/4/2001 1:24:38 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2947	BOD	Unknown Time Adjustment - Neph_Calib
10/4/2001 1:24:38 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2947	BOD	Unknown Precision (values left unrounded) - Neph_Calib
10/4/2001 1:24:37 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2947	BOD	Unknown Time Adjustment - Neph_RH
10/4/2001 1:24:37 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2947	BOD	Unknown Precision (values left unrounded) - Neph_RH
10/4/2001 1:24:37 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2946	BOD	Unknown Parameter Code - Neph_RH
10/4/2001 1:24:37 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2947	BOD	Unknown Averaging Period - Neph_RH
10/4/2001 1:24:37 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2946	BOD	Unknown Parameter Code - Neph_Lamp
10/4/2001 1:24:37 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2946	BOD	Unknown Parameter Code - Neph_Calib
10/4/2001 1:24:37 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	2946	BOD	Unknown Parameter Code - Neph_Scat1
10/4/2001 1:24:36 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1485	BOD	Unknown Time Adjustment - Neph_Lamp
10/4/2001 1:24:36 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1485	BOD	Unknown Precision (values left unrounded) - Neph_Scat1
10/4/2001 1:24:36 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1485	BOD	Unknown Time Adjustment - Neph_Scat1
10/4/2001 1:24:36 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1485	BOD	Unknown Averaging Period - Neph_Calib
10/4/2001 1:24:36 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1485	BOD	Unknown Precision (values left unrounded) - Neph_Calib
10/4/2001 1:24:36 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1485	BOD	Unknown Time Adjustment - Neph_Calib
10/4/2001 1:24:36 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1485	BOD	Unknown Averaging Period - Neph_Lamp
10/4/2001 1:24:36 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1485	BOD	Unknown Precision (values left unrounded) - Neph_Lamp
10/4/2001 1:24:35 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1485	BOD	Unknown Averaging Period - Neph_RH
10/4/2001 1:24:35 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1484	BOD	Unknown Parameter Code - Neph_RH
10/4/2001 1:24:35 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1484	BOD	Unknown Parameter Code - Neph_Scat1
10/4/2001 1:24:35 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1484	BOD	Unknown Parameter Code - Neph_Lamp
10/4/2001 1:24:35 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1485	BOD	Unknown Precision (values left unrounded) - Neph_RH
10/4/2001 1:24:35 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1485	BOD	Unknown Time Adjustment - Neph_RH
10/4/2001 1:24:35 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1485	BOD	Unknown Averaging Period - Neph_Scat1
10/4/2001 1:24:35 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	1484	BOD	Unknown Parameter Code - Neph_Calib
10/4/2001 1:24:33 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	23	BOD	Unknown Time Adjustment - Neph_RH
10/4/2001 1:24:33 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	23	BOD	Unknown Precision (values left unrounded) - Neph_Lamp
10/4/2001 1:24:33 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	23	BOD	Unknown Time Adjustment - Neph_Calib
10/4/2001 1:24:33 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	23	BOD	Unknown Averaging Period - Neph_Calib
10/4/2001 1:24:33 PM	\\Netserver2\CRPAQS\DataMgt\ftpRecent\bod recent.dat	23	BOD	Unknown Time Adjustment - Neph_Scat1

Record: 7 of 371819
Datasheet View NUM

Figure 2-7. Excerpt from LogImport table in the database.

2.3.3 Handling Missing Data

Missing data appeared in the Recent.DAT file marked as "M" except in cases where the DAS went off-line and the only indication of missing data were gaps in the DAS time. The DAS produced a record for all instruments in the Recent.DAT file every minute, regardless of individual instrument sampling frequencies. For parameters with sampling intervals greater than the 1-minute resolution of the DAS, "M" was assigned as the data value for those records between sampling records although they were not actually "missing." Upon import, the CRPAQS data management system did not translate missing data or data marked with an "M" because it was unable to discern between actual missing records and dummy "M" records associated with parameters that had sampling frequencies greater than 1 minute.

Data gaps needed to be filled in the database to facilitate later averaging of the data (e.g., computing 60-minute averages from 1-minute data). Missing data records were added to the database by manually running a program written specifically to fill in missing records. This program compared the actual time interval between records to the expected time interval and added missing records to the database with a Data Value = -99 and QC code = 9 for the cases where the time difference between two subsequent records was greater than the sampling interval.

2.3.4 Other Database Issues

As the study progressed, especially during the winter intensive when a large volume of data was being processed daily, the burgeoning 1-minute database became extremely unwieldy and caused numerous problems. Slow processing, resulting from database space issues, was initially mitigated by splitting the database into two parts; however, this solution was only temporary. Eventually, following a server upgrade, the databases were successfully recombined and proved reliable. The 5-minute database was also affected by the large influx of data (and temporarily split into two as well), but the comparatively smaller 10-minute and 60-minute databases were generally stable throughout the course of the study. Occasional server reboots, reestablishment of ODBC links between database and server, and the compaction and optimization of the databases were required.

2.4 STI'S DAILY DATA QC

The purpose of the daily review of data at the DMC during the field study was to ensure proper operation of the field equipment and to provide feedback to the field operators when potential problems were identified. Either the field operators or the field manager also performed daily reviews of the data, checked instruments, prepared daily site reports, and corrected instrument problems promptly. Daily operations at STI were typically performed Monday through Friday.

2.4.1 Review of Daily Plots of the Data

Once the data were received at the DMC each morning, several automatic functions designed to prepare the data set for initial review were performed. The initial review was designed to quickly identify both stable and proper operation and unstable or improper operation. Each morning, two-day running time series plots were automatically prepared for selected parameters. In addition, time series plots were automatically prepared for the daily zero and span results. Each morning, a scientist reviewed these time series plots and identified potential problems or issues to be resolved. If a problem was identified, it was discussed with a senior scientist who reviewed the proposed solutions and recommended the approach to be used. Communication with the field operations personnel was prompt. When the solution to the problem was not quickly identifiable (and it was not resolved), or if the problem was persistent, it was discussed with the instrument expert and a more detailed analysis was performed.

Every morning during the field study, two sets of data plots were automatically printed (see **Table 2-7** for a list of plots). Several actions resulting from these plots were performed by an STI data technician. The first set of plots showed a two-day time series of selected parameters (see **Figure 2-8**). For example, the plots printed on the morning of March 14, 2000, show data from midnight March 12 to midnight on March 14. The following parameters were plotted:

- Gaseous species: ozone, nitrogen oxide/oxides of nitrogen (NO/NO_y), nitric acid (HNO_3), peroxyacetylnitrate (PAN), nitrogen dioxide (NO_2), sulfur dioxide (SO_2)
- Particulate matter: nephelometer particle scatter (b_{sp}) including relative humidity (RH) (both analog and digital signals); BAM PM_{10}, BAM $PM_{2.5}$ mass concentration and volume; Climet optical particle counter (OPC) sample flow and particle count, Particle Measuring Systems (PMS) Lasair OPC particle count and volume, scanning mobility particle counter (SMPS) particle count and sample flow; $PM_{2.5}$ sulfate concentration; $PM_{2.5}$ nitrate concentration
- Carbon: organic carbon and elemental carbon (OCEC) total, OC concentrations and flow; 1-wavelength (880 nm) Aethalometer black carbon (BC) and flow; 7-wavelength Aethalometer BC at ultraviolet (UV) 350, Blue 450, Green 571, Yellow 590, Red 660, near infrared (NIR) 880, NIR 950 and flow

The second set of plots contained calibration information for the current morning (e.g., **Figure 2-9**). The calibration data were plotted for NO/NO_y, ozone, nitric acid, PAN/NO_2, and SO_2 instruments. The time scale was expanded so that the analyst could better examine the zero, span, and calibration information. Calibration plots are discussed in Section 3.

In addition to printed copies of the plots, the plots were also automatically posted in *.GIF format to a special STI data web site for convenient viewing. The daily time series plots posted on the Internet were extremely useful to the field staff, data management staff, and measurement experts. **Figure 2-10** shows an example of an instrument malfunction. Using the plots, the instrument problems were identified quickly and corrected, thus minimizing data losses.

Table 2-7. Parameters for which daily time series plots were prepared (by site).

Site	Site Code	Parameter								
		OCEC	BC	BAM $PM_{2.5}$	b_{sp}	NO, NO_y	Ozone	OPC[b]	SO_2	Other
Altamont	ALT1			✓						
Angiola[a]	ANGI	✓	✓	✓	✓	✓	✓	✓		Nitric acid, nitrate, PAN, NO_2
Bakersfield	BAC	✓	✓	✓	✓	✓	✓	✓	✓	Nitrate, PAN, NO_2
Bodega	BODB		✓							
Bethel Island	BTI		✓	✓	✓	✓				Nitrate, PAN, NO_2
Corcoran	COP		✓	✓	✓					Nitrate
Edwards AFB	EDW		✓	✓	✓					
Fresno Supersite	FSF	✓	✓	✓	✓	✓	✓	✓		
Modesto, 14th Street	M14		✓		✓					
Sacramento Del Paso Manor	SDP		✓	✓	✓					
San Jose	SJ4		✓	✓	✓					Nitrate
Sierra Nevada Foothills	SNFH		✓	✓	✓	✓	✓			Nitrate, PAN, NO_2, meteorology
Walnut Grove	WAG		✓		✓					Nitrate
Walnut Grove Tower	WGT		✓		✓					Nitrate

[a] Plots were also made of Angiola Tower instrument measurements.
[b] Typically included Climet, PMS, and SMPS measurements.

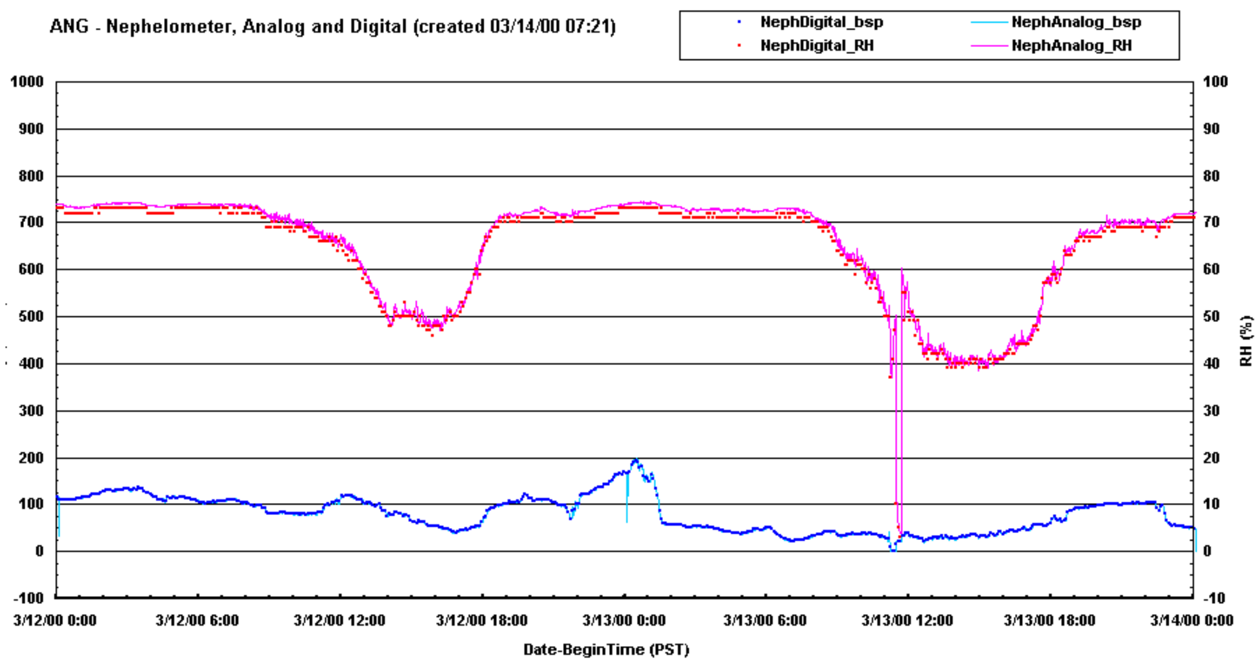
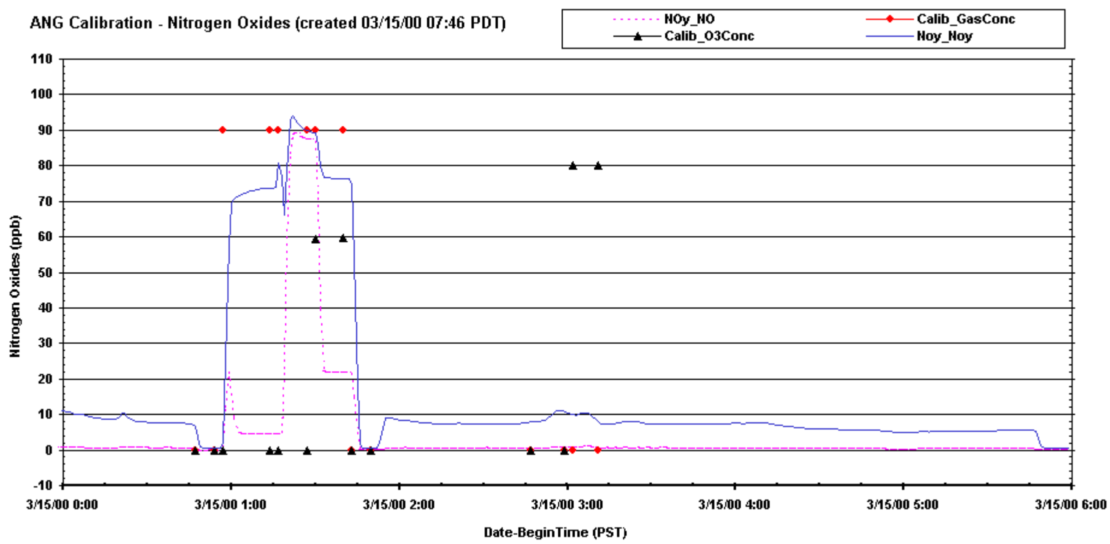


Figure 2-8. Example two-day time series plot for nephelometer data posted to the STI CRPAQS web site for daily QC purposes.



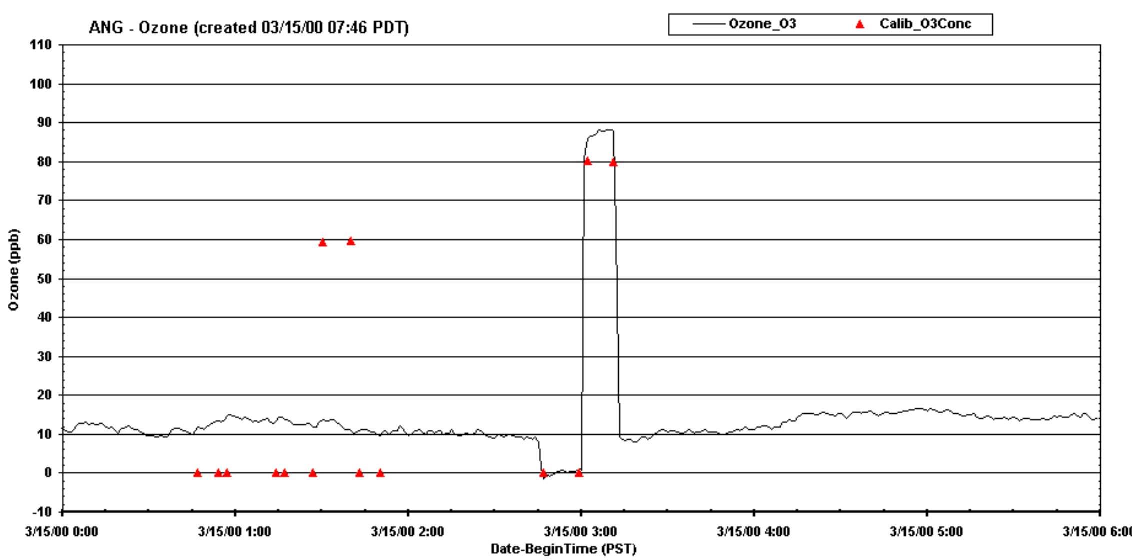


Figure 2-9. Example calibration plots posted to the STI CRPAQS web site for daily QC purposes.

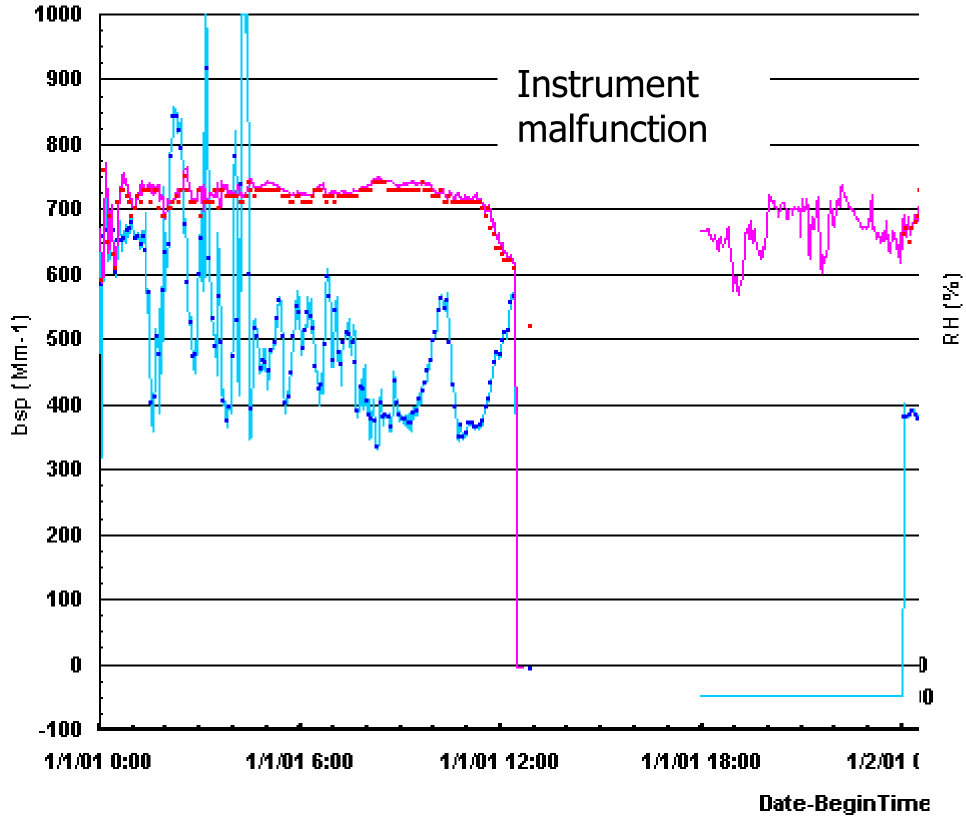


Figure 2-10. Example of an instrument malfunction (Angiola nephelometer) illustrating the usefulness of the daily time series plots posted on the STI CRPAQS web site.

Monday through Friday during the field study, an STI data technician collected the plots from the printer and performed the following tasks:

- Faxed a copy of all plots to the Bakersfield office for the field manager.
- Faxed a copy of the Angiola plots to the Angiola site.
- Three-hole punched the printed plots and physically posted them at STI in a location dedicated to CRPAQS data. Plots were arranged by site and measurement type. These plots were archived after the study.

The STI data technician reviewed either the on-line or printed plots for missing data, odd data, and other problems. If problems were noted, the data technician contacted the appropriate field managers and site technicians by phone and e-mail to alert them. At this stage, the site logs (faxed daily to STI from the field) were useful in assessing what actions were taken on the sampling equipment on the previous day. For example, site log files were used to identify when an instrument was taken off-line for maintenance.

The STI data technician recorded observations about the data in a Microsoft (MS) Excel file. An excerpt from this MS Excel file is provided in **Table 2-8**. This table illustrates the running compilation of observations and explanations of problems obtained from site or

Table 2-8. Example comments made during daily data QC. Explanations usually came from daily site or instrument logs and/or discussions with site personnel.

Site	Date	Start	End	Instrument	Comment	Explanation
ANG	27-Feb	1900	1900	BAM	PM2.5 > PM10	
ANG	27-Feb	800	2400	BAM	Low concs.	
ANG	27-Feb	all	all	Aeth	Low flow (6.5 LPM), BC low	
ANG	26-Feb	all	all	OCEC	flow = 0, OC and OCEC non -zero	
ANG	27-Feb	all	all	OCEC	flow = 0, OC and OCEC non -zero	
BAC	27-Feb	100	200	NO/NOy	response to cal good	
BAC	27-Feb	350	350	NO/NOy	NO > NOy	
BAC	27-Feb	400	1000	NO/NOy	spikes	
BAC	27-Feb	0	0	Neph	dip in analogue bsp	
BAC	27-Feb	900	1700	BAM	Negative PM25, corresponding low PM10	
SDP	27-Feb	0	0	Neph	dip in analogue bsp	
SDP	27-Feb	0	2400	Aeth	Low flow (6.0 LPM)	
SDP	28-Feb	0	0	Neph	dip in analogue bsp	
SDP	28-Feb	0	2400	Aeth	Low flow (6.0 LPM)	
ANG	28-Feb	1700	2300	all	data gap	DAS reset (clear). Reset at 8 a.m. 2/29/00
ANG	28-Feb	300	300	Ozone	span = 85, O3 response = 90	
ANG	28-Feb	0	0	Neph	dip in analogue bsp	
ANG	28-Feb	0	1000	Neph	RH > 80	Worked on temp controller at 1400
ANG	28-Feb	1300	1300	Neph	dip in RH, slight bsp increase	Calibrated at 1235-1415. Worked on temp controller at 1400
ANG	28-Feb	0	2400	BAM	PM2.5 > PM10 intermittent (200 and 300 significant)	
ANG	28-Feb	0	2400	BAM	BAM10 and BAM25 values reported at slightly different times	
ANG	28-Feb	1600	2400	BAM	data gap	
ANG	28-Feb	1500	2400	OCEC	data gap	
ANG	28-Feb	930	1630	Aeth	Oscillating around zero	
ANG	27-Feb	950	1000	NO/NOy	noise seen by Chuck	capillaries cleaned due to alarm.
BAC	28-Feb	2130	2130	NO/NOy	NOy spike	seen at site DAS
BAC	28-Feb	0	0	Neph	dip in analogue bsp	
BAC	28-Feb	1430	1430	Neph	dip in RH, slight bsp increase	zero and span performed between 1418 and 1455
BAC	28-Feb	100	600	BAM	Negative PM25, corresponding low PM10	
BAC	28-Feb	1600	1700	BAM	Negative PM25	
BAC	28-Feb	1700	2200	BAM	Deviation in BAM10 and/or BAM25 flow	BAM10 filter tape changed at 1615 to 1630
BAC	28-Feb	1900	2400	Aeth	significant increase in BC concs (and flow)	Corresponds with increase in PM2.5, bsp, and NO/NOy
BAC	28-Feb	100	200	NO/NOy	response to cal good, cal gas = 90, response 87	
ANG	29-Feb	0	800	all	data gap	
ANG	29-Feb	0	800	Neph	gap in digital data?	
ANG	29-Feb	0	700	BAM	data gap (shorter?)	
ANG	29-Feb	1300	1300	BAM	dip in Bam10Vol, peak in PM10 mass conc.	
ANG	29-Feb	1400	1600	BAM	PM2.5 > PM10	
ANG	29-Feb	2000	2200	BAM	PM2.5 > PM10	
ANG	29-Feb	0	700	OCEC	data gap (shorter?)	
BAC	29-Feb	30	1100	NO/NOy	extreme NOy spikes	
BAC	29-Feb	1830	1830	NO/NOy	NO spikes later than NOy	Instrument worked on between 1430 and 1745
BAC	29-Feb	915	915	NO/NOy	NO > NOy	

equipment log sheets and through discussions with site personnel. These observations were referenced in later data validation efforts. Common problems and features that the STI data management technician looked for in the data are listed in **Table 2-9**.

Table 2-9. Common problems and features in the data for which the technician was instructed to look. Abbreviations are defined at the beginning of this report.

Instrument/Parameter	Feature/Problem	Action/Explanation
All	Missing data Negative concentrations Quick drop to zero Large spike in concentration	Alert appropriate person
Aethalometer	Periodic dip in flow rate and b_{sp}	Symptom of a tape transfer
Aethalometer	Flow rate > 7 or < 6.5 LPM	Note in log
Aethalometer	Extreme noise	Note in log
BAM	$PM_{2.5} > PM_{10}$	Note in log
BAM	Constant data values	Tape problem
Nephelometer	Dip or spike in RH and b_{sp} at 0000	Corresponds to GetNeph program execution
Nephelometer: RH	RH > 75%	Heater problem

Table 2-9. Common problems and features in the data for which the technician was instructed to look. Abbreviations are defined at the beginning of this report.

Page 2 of 2

Instrument/Parameter	Feature/Problem	Action/Explanation
Nephelometer	Difference between analog and digital signals	Interference between both signals or problems related to one signal
Nephelometer	Flat signal or no signal in either or both channels	GetNeph or DAS glitch; required auto restart at 0000
Nephelometer	Extreme noise	Not commonly observed
Nitrate and Sulfate	Missing data	Possible flash strip problem
Nitric Acid	$[NO_y\text{-}HNO_3]>[NO_y]$	Note in log
NO/NO_y	Low concentrations	If correspond with higher ozone, may be real
NO/NO_y	High concentrations, noise in afternoon in SNFH data	Daily fire burns east of the site, starting approximately 1500 PST daily
NO/NO_y	$NO>NO_y$	Note in log
PAN/NO_2	Extreme noise	Note in log
OCEC	OC>OCEC	Note in log
Ozone	Low concentration; dips in data	If corresponds with higher NO/NO_y, may indicate real titration

2.4.2 Calibration, Zero, and Span Information

During the field study, the STI data technician performed two tasks to QC the calibration, zero, and span information. First, the technician reviewed the current day's zero and span data plotted for the Angiola, Bakersfield, Sierra Nevada Foothills, and Bethel Island sites. The calibration gas concentration was compared to the measured concentration, and differences greater than 10% were noted and communicated to the field manager. Second, the technician compared the actual time of the calibration to the scheduled time. Missing calibrations were noted and communicated to the field manager. Calibration plots were compared over a period of several days to determine whether instrument response had drifted. Daily comments about calibration problems were added to the MS Excel comments file for future reference.

Additional discussion of the calibration, zero and span information is provided for individual instruments in later sections of this report.

2.5 PARAMETER PROCESSING

Each parameter/instrument combination has unique properties with respect to its sampling frequency, sampling time, detection limit, calibration schedule, et cetera. We have graphically summarized the flow of data for each instrument to accompany the text in the following sections.

2.5.1 Ozone

Figure 2-11 illustrates the flow of data for the ozone instrument, from sampling to storage in the SQL server. The flow diagram provides details of the analyzer (measurement cycle, averaging period, detection limit, analytical precision), DAS information (time stamp offset, stored data precision, file naming convention), database files (directory names), maintenance (checks, filter changes), multipoint calibrations (frequency, concentrations), and zeroes and spans.

The API Model 400 Ozone Analyzer provided 1-minute ozone concentration (ppb) data for the Angiola, Bakersfield, and Sierra Nevada Foothills sites. Ozone data were imported and stored in the 1-minute SQL database. Reports and plots were generated for the Bakersfield site to assist other investigators, but the data were not evaluated or QC'd by STI. Ozone and NO/NO_y data were plotted together in the reports in order to monitor the relationship between the gaseous species. Nightly zero-span calibration data were sent after 0600 PST every morning to be plotted (on an expanded 6-hr time scale) along with the two-day running plots.

There were relatively few problems processing the ozone data. At Angiola, the instrument clock ran slower than the DAS clock from March through May 2000, resulting in intermittent data gaps in the Recent.DAT file and, subsequently, the database.

2.5.2 NO/NO_y

NO and NO_y concentration (ppb) data from the Thermo Environmental 42 CY NO_y instrument were imported into and stored in the 1-minute database for the Angiola, Bakersfield, Sierra Nevada Foothills, and Bethel Island sites. **Figure 2-12** shows the flow of data for this instrument and provides details of the analyzer (data collected, averaging period, detection limit), DAS information (time stamp offset, stored data precision, file naming convention), database files (directory names), maintenance (checks, filter changes), calibrations (frequency, concentrations), and zeroes and spans.

Data included nightly zero span checks and five daily matrix zero checks flagged with the appropriate OP code. Both two-day running plots and nightly zero span plots (on an expanded 6-hr time scale) were created each morning. Most problems related to the import and processing of the NO/NO_y data were general to the 1-minute database and were discussed previously.

A unique feature of the operation of the instruments was the use of instrument-specific codes to signify key events. For the NO/NO_y instrument, for example, unique OP codes indicated when each of the span and calibration gases were run through the instrument (including NO, NO_2, ammonia, n-propylnitrate [NPN], and nitric acid). **Figure 2-13** shows an example of a NO_y calibration plot for Angiola on August 3, 2000.

2.5.3 Nephelometer

Figure 2-14 shows the data flow for the Radiance Research M903 Integrating Nephelometer, and provides information about the analyzer (data collected, averaging periods, accuracy), data acquisition (via DAS or GetNeph program), data storage, and calibration and zero span checks. This instrument collected the light extinction coefficient from scattering by

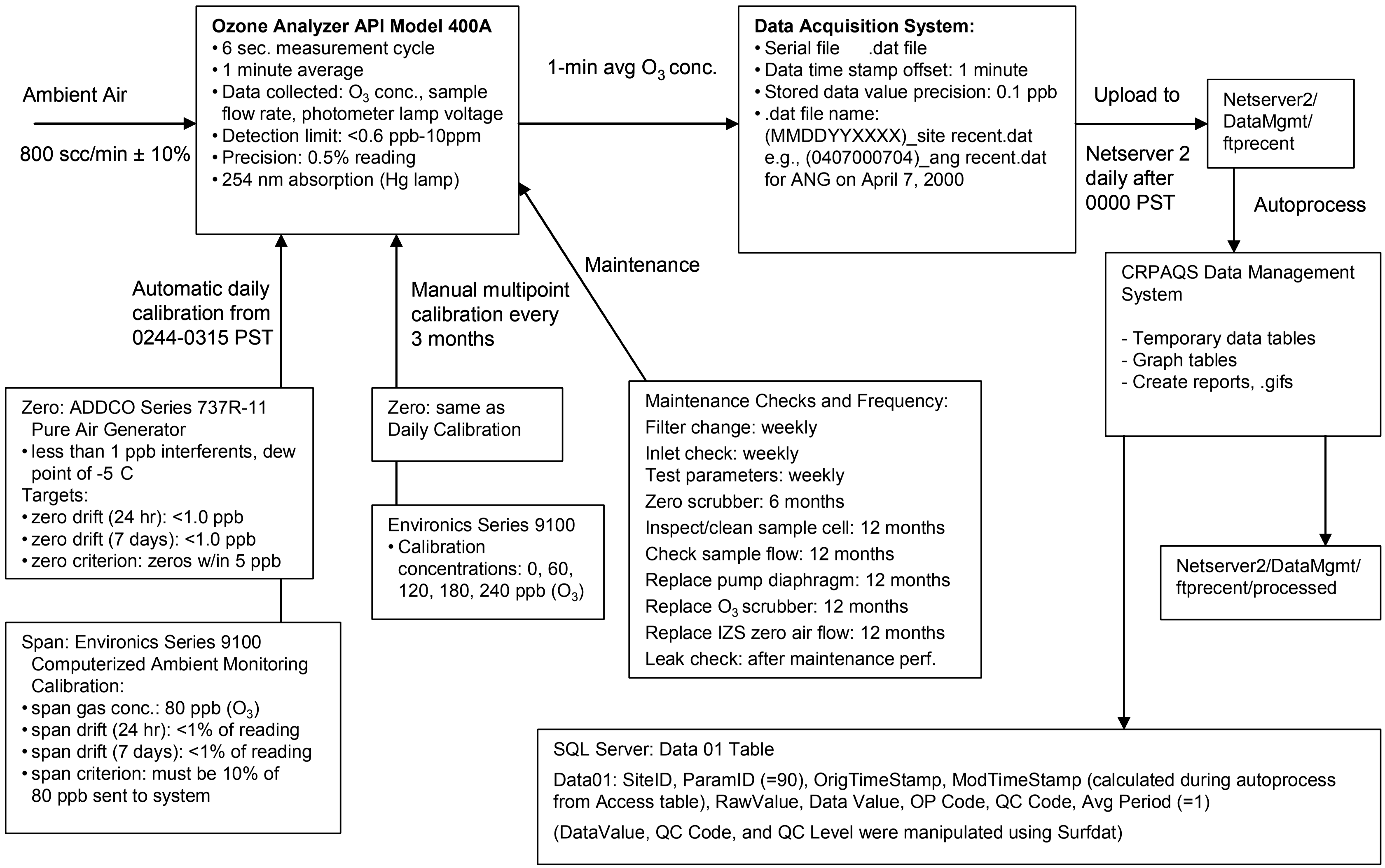


Figure 2-11. Flow chart for ozone data.

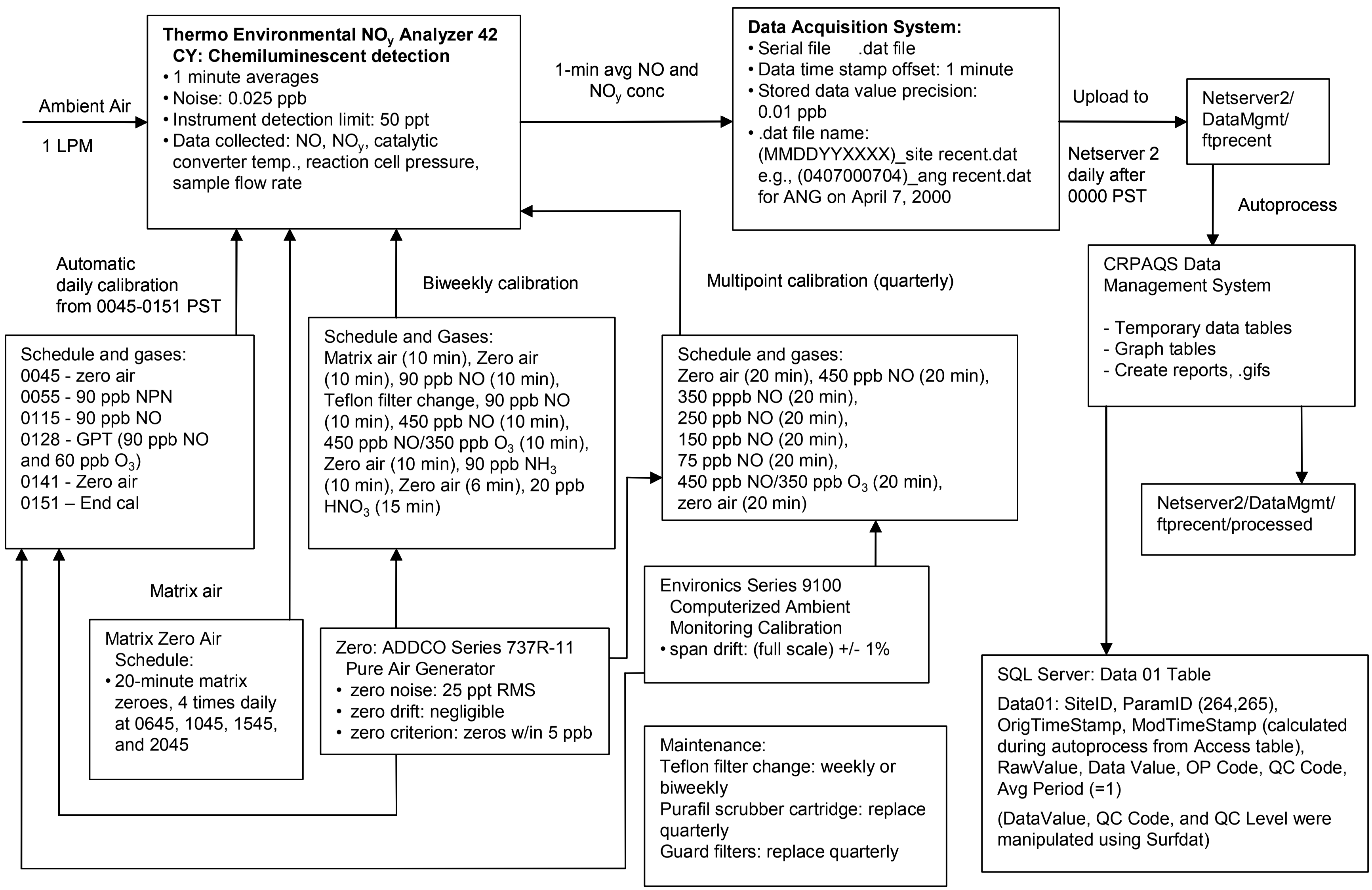


Figure 2-12. Flow chart for NO/NO$_y$ data.

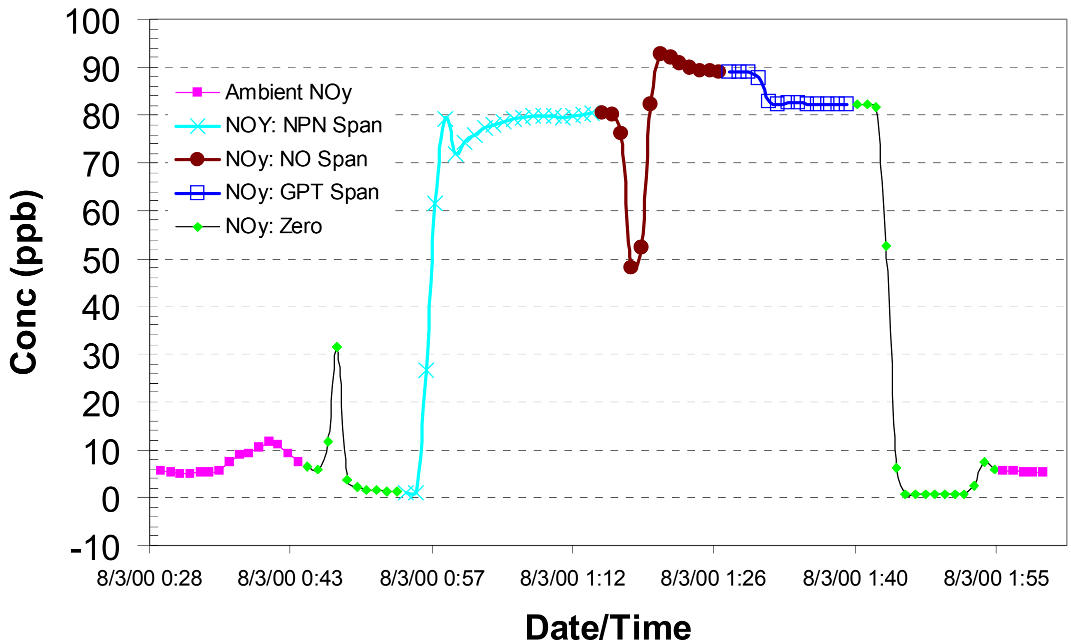


Figure 2-13. Illustration of instrument-specific OP codes used to signify key events including calibrations. This example for the NO/NO_y instrument shows how OP codes were used to indicate when each of the span and calibration gases were run through the instrument. NPN is n-propylnitrate and GPT is gas-phase titration.

particles (b_{sp}), relative humidity (RH), temperature (T), pressure (P), calibration signal, and lamp voltage data that were subsequently imported and stored in two different databases, 1-minute and 5-minute, according to the sampling interval of the data record. Analog b_{sp}, RH, lamp voltage, and calibration signal data were stored in the 1-minute table, and digital b_{sp}, RH, T, P, and calibration signal data were stored in the 5-minute table. Three nephelometers (on the Angiola Tower) also measured and reported P.

The *.NPH file, containing the 5-minute data, was generated daily by the GetNeph program, which ran automatically after midnight and retrieved digital data from the instrument. One-minute analog data were written to the *.DAT file on a continuous basis. STI received both analog and digital data from the Angiola, Bakersfield, Walnut Grove Tower, San Jose, and Sacramento sites; digital data only from the Corcoran and Angiola Tower sites; and analog data only from the Bethel Island, Sierra Nevada Foothills, and Walnut Grove sites. The digital data in the *.NPH files were directed through a different parser and processed separately from the analog data in the *.DAT files. For the sites where both analog and digital data were available, the data were plotted together in daily QC graphs. Nephelometer and BAM data were printed on the same page. Some of the data at Sacramento were received on disk from T&B Systems.

The 5-minute digital data were used exclusively in the study, except in cases where only analog 1-minute data were available, because the digital reading is a more accurate assessment. One-minute data were collected to ensure a continuous measurement. At sites like Sacramento, where the digital channel was disconnected partway into the study due to interference issues, the 1-minute analog data were averaged and delivered in lieu of the 5-minute digital data.

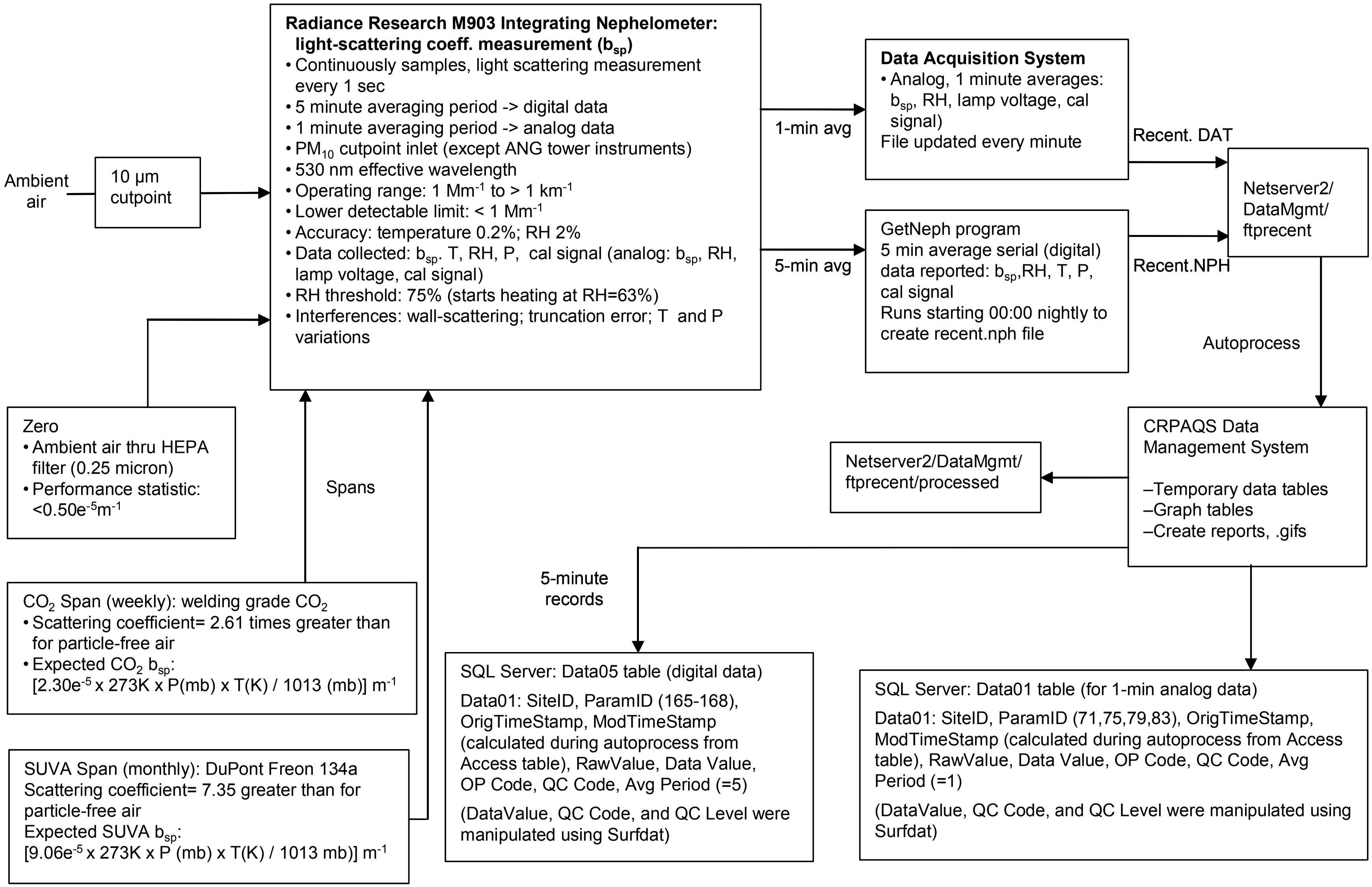


Figure 2-14. Flow chart for nephelometer data.

2.5.4 BAM $PM_{2.5}$ and PM_{10}

The Met One 1020 Beta Attenuation Mass Monitor (BAM) provided information on particle mass concentration ($\mu g/m^3$) and volume (m^3) of particles entering a size-selective inlet of either 2.5 or 10 microns. **Figure 2-15** shows the data flow for this instrument. Analyzer information (measurement cycle, stability, accuracy, etc.), flow calibration/audit schedules, and data acquisition and data storage issues (file naming conventions, data record structure) are outlined in the diagram. For all sites except Altamont, the 1-hr mass concentration and volume data (contained within a Recent.DAT file) were processed and imported nightly to the 60-minute database. BAM $PM_{2.5}$ data from Altamont were contained within a serial file (*.BAM), manually dumped in the STI database once a month, and processed using a parser specific to the BAM data file format. When both sets were available, BAM $PM_{2.5}$ and BAM PM_{10} data were plotted on the same graph.

No major problems were associated with the import and processing of the BAM data; however, some minor difficulties required attention. The automatic translation of the BAM instrument error code (OP code = 7) to an invalid QC code (8) was discontinued after it was discovered the instrument gave error codes (relating to "deviant membrane density") due to a membrane modification. At Angiola, the serial cable became disconnected from the PM_{10} instrument on January 18-19, 2001, affecting the proper time stamps on the records. These offset times were manually readjusted in the database. The $PM_{2.5}$ and PM_{10} cyclones on the BAM instruments at Angiola were switched early in the study, requiring a redefinition of the instruments and edit of the initialization file on the DAS, but processing of the data was unaffected.

2.5.5 Aethalometers

Two Aethalometer models were used in this study: the Andersen Instruments 1-wavelength (AE-1X) and the 7-wavelength (AE-3X) Aethalometers. Data flow diagrams for the two instruments are provided in **Figures 2-16 and 2-17**, showing information about the analyzer (sensitivity, interferences, tape mechanism issues, etc.), data acquisition and data storage (file names, data record structure), and various checks (weekly, monthly, and bimonthly). Parameters collected from the Aethalometer included BC (880 nm) concentration and flow for the 1-wavelength instrument; and BC at 370 (UV), 470 (blue), 520 (green), 590 (yellow), 660 (red), 880 (NIR), and 950 (NIR) nm and flow for the 7-wavelength instrument. The 7-wavelength Aethalometer is designed so that the concentration derived from absorption for all wavelengths is identical unless the aerosol is active at a particular wavelength (e.g., higher UV absorbance could indicate the presence of wood-burning smoke (Magee Scientific Company, 2002)). The parameters ingested in each record comprise each wavelength's absorption converted into concentrations, as well as the instrument's flow rate.

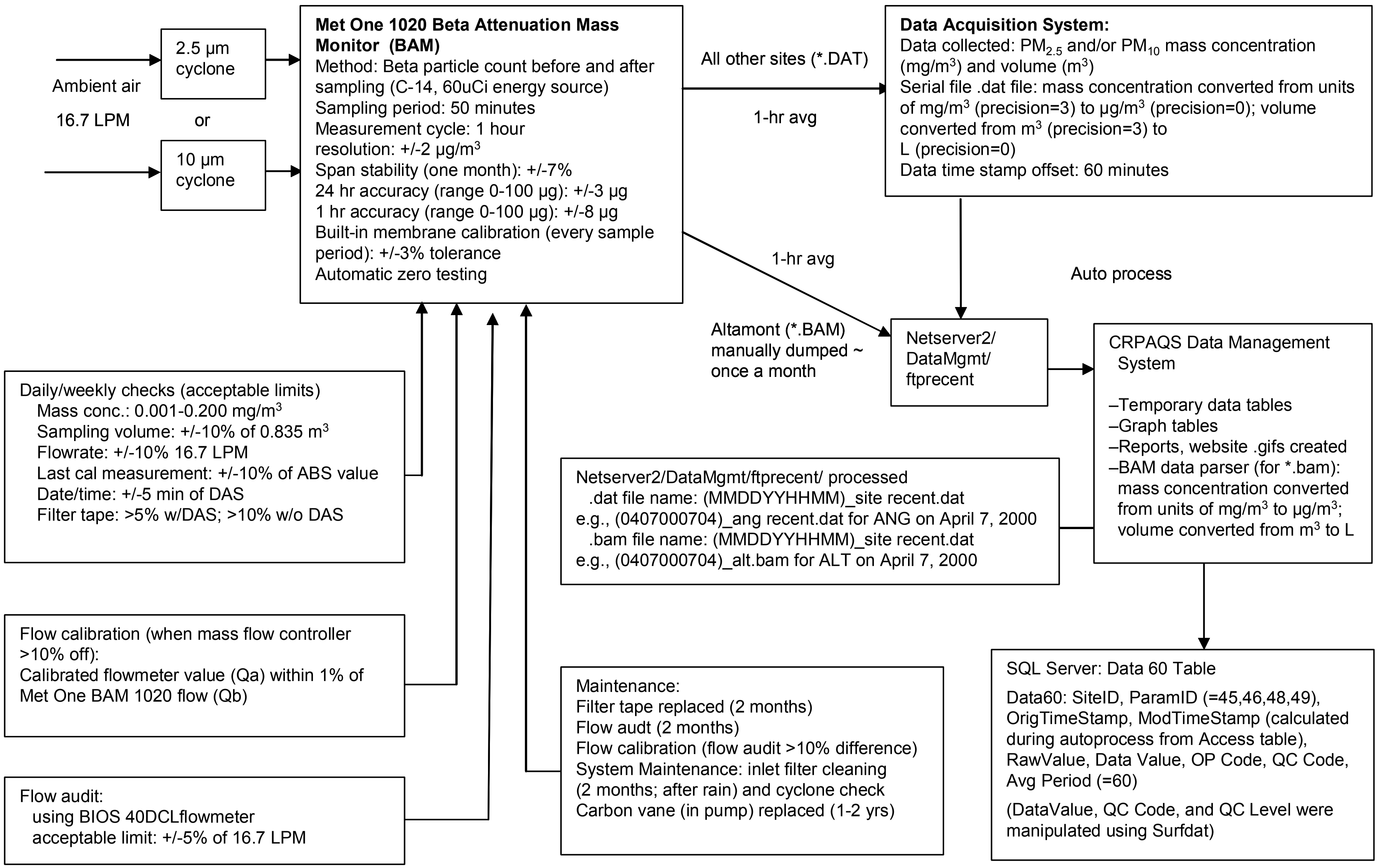


Figure 2-15. Flow chart for BAM $PM_{2.5}$ and BAM PM_{10} data.

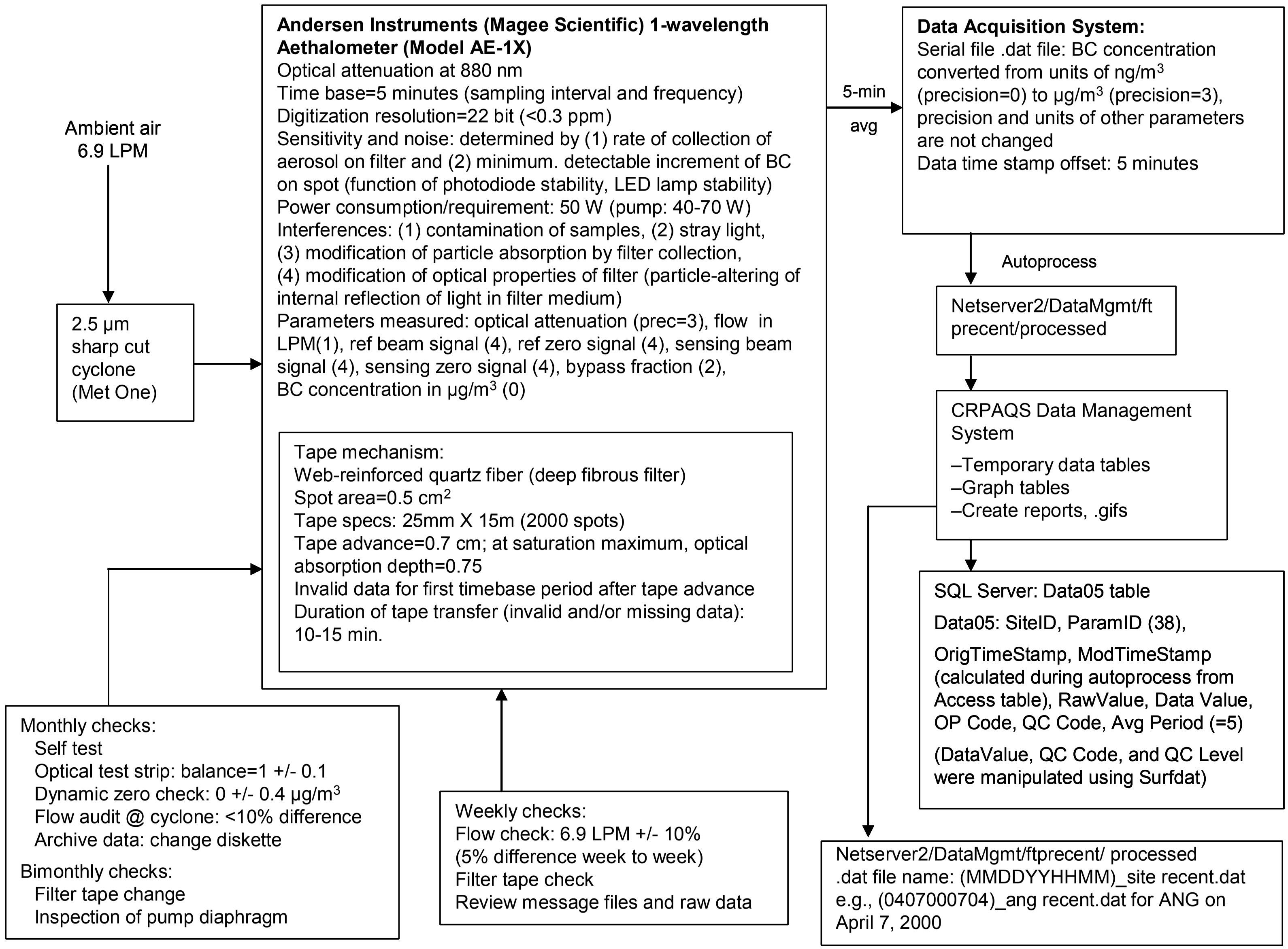


Figure 2-16. Flow chart for 1-wavelength Aethalometer data.

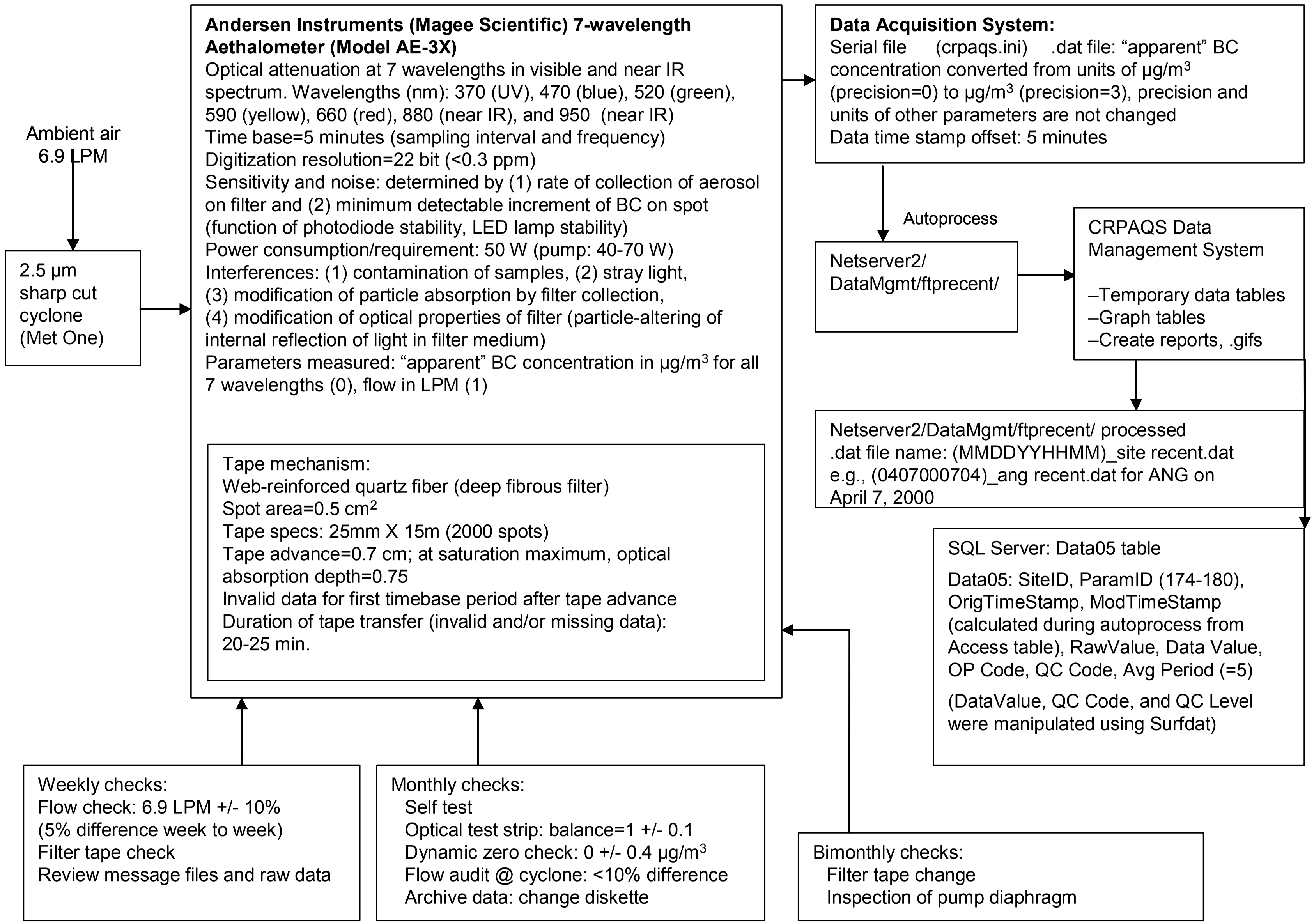


Figure 2-17. Flow chart for 7-wavelength Aethalometer data.

One-wavelength Aethalometers were installed at several sites between January and February 2000. Starting in November 2000, a number of sites (Angiola, Bakersfield, Sacramento, and San Jose) substituted the 7-wavelength instrument for the 1-wavelength instrument, requiring necessary changes in the database to facilitate the new parameters at the affected sites. The flow parameter for the 7-wavelength instrument was initially assigned to the same parameter as the 1-wavelength instrument. However, prior to data validation, the flow for the 7-wavelength instrument was given a unique parameter identification (to differentiate the data from the two instruments), and the data records were updated to reflect this change. During the first few days of collection from the Walnut Grove and Walnut Grove Tower sites, Aethalometer data were assigned an incorrect parameter name; these data were reprocessed after the names were corrected.

The largest problem in recovering and processing the Aethalometer data concerned incorrectly averaged data values recorded during instrument tape transfer. An Aethalometer tape transfer occurred when the aerosol BC surface loading reached 4 $\mu g/cm^2$, or an optical depth of 0.75. The spot surface area is 0.5 cm^2. For the 1-wavelength instrument, the mechanical transfer lasted 10 to 15 minutes, generally resulting in the absence of two to three consecutive data records. For the 7-wavelength instrument, the mechanical transfer lasted 20 minutes, generally resulting in the absence of four consecutive data records. Revised 7-wavelength Aethalometer software (v .093f), which was installed any time on or after December 18, 2000, effected a shortened tape advance of fifteen minutes and the loss of only three records. Before, during, and after the tape advances, data values were missing, invalid, or, in some cases, incorrectly averaged but usable by application of a correction factor. As noted above, at the start of the tape transfer, the instrument produced two to five successive null values, some occurring within the same minute as the last valid data value. The valid and null data values were averaged together into one record by the DAS, which was operated on a 1-minute time resolution. These truncated data records occurring during the tape transfer were queried from the database, and an appropriate correction factor, based on the total number of values averaged, was applied. The corrected records were then reingested into the database before data validation commenced.

The length of tape transfer and total data loss were dependent on the combination of instrument model, use/disuse of the tape-saver function, DAS software version, and emission source activity near the sampling site (i.e., particle loading on the tape). The following provides a general guide to data loss:

- 1-wavelength instrument: 1 to 3 tape advances a day, 15 minutes each (15 to 45 minutes a day total data loss)
- 7-wavelength instrument, no tape saver, old software: 10 to 20 tape advances a day, 25 minutes each (250 to 500 minutes a day total data loss)
- 7-wavelength instrument, 10X tape saver, old software: 5 to 10 tape advances a day, 25 minutes each (125 to 250 minutes a day total data loss)
- 7-wavelength instrument, 10X tape saver, new software: 1 to 7 tape advances a day, 20 minutes each (20 to 140 minutes a day total data loss)

Table 2-10 provides a summary of the software changes for the Aethalometers.

Table 2-10. Summary of software changes for the Aethalometer; changes reduced data loss.

Site	Serial Number (S/N)	Date/Time of Change from Old to New Software
SDP	271	1/17/2001 1425
SJ4	273	1/19/2001 1541
EDW	(256)[a]	n/a
SNFH	272	1/16/2001 1405
COP	(271)[a]	n/a
BTI	255	1/18/2001 1459
BAC	256	1/8/2001 1250
ANGI	257	1/17/2001 0000
ANG100	254	1/18/2001 1459

[a] These instruments were no longer deployed by the time the software change was effected.

One significant pattern that should be considered by data users was observed. Immediately following tape advances, individual data were often 20% to 40% higher than real values ((Hansen, 2002; LaRosa et al., 2002). This enhancement, which often persisted from 5 to 30 minutes (one to six records) after a tape advance, is due to the nature of the first-deposited layer on the quartz filter tape, which has small, electrophilic cracks. When particles of varying composition deposit in these cracks, the light is refracted; the particle's microenvironment results in multiplicative optical enhancement. Because this phenomenon is thought to be universal in Aethalometer data, these apparently high data were not flagged as suspect. Indeed, there is even evidence to support that the few data points preceding a tape advance are lower than real values and thus, over longer averaging periods, "compensate" somewhat for the high values succeeding tape advances (LaRosa et al., 2002).

2.5.6 OCEC

Sixty-minute data collected from the Rupprecht and Patashnick Series 5400 Ambient Carbon Particulate Monitor included OC and OC plus EC (OCEC) mass concentration and volume (flow), as well as the instrument's afterburner temperatures. A schematic of the data flow for the continuous OCEC instrument is shown in **Figure 2-18** and details the monitor (method, instrument cycles, data collected), data acquisition (time stamp issues, file naming conventions), data storage (structure of the data record, storage tables), audit/calibration information (types of audits and frequency), and maintenance tasks (cleaning, replacing filters, etc.). Flow data in liters per minute (LPM) were collected until April 19, 2000, when the parameter was replaced by related volume (L) data, the gathering of which continued through the end of the study.

The first instruments at the Angiola and Bakersfield sites presented various operational problems, resulting in weeks of missing data (see Wittig et al., 2003) instruments at both sites were eventually replaced and fewer problems were encountered. Although the sampling interval for OCEC data was 60 minutes, the start time (ModTimeStamp) for each data record was rolled back 120 minutes from the DAS time to account for the 1-hr analysis period following the 1-hr sample collection. Some operating parameters were not wholly consistent throughout the field study (see **Table 2-11**).

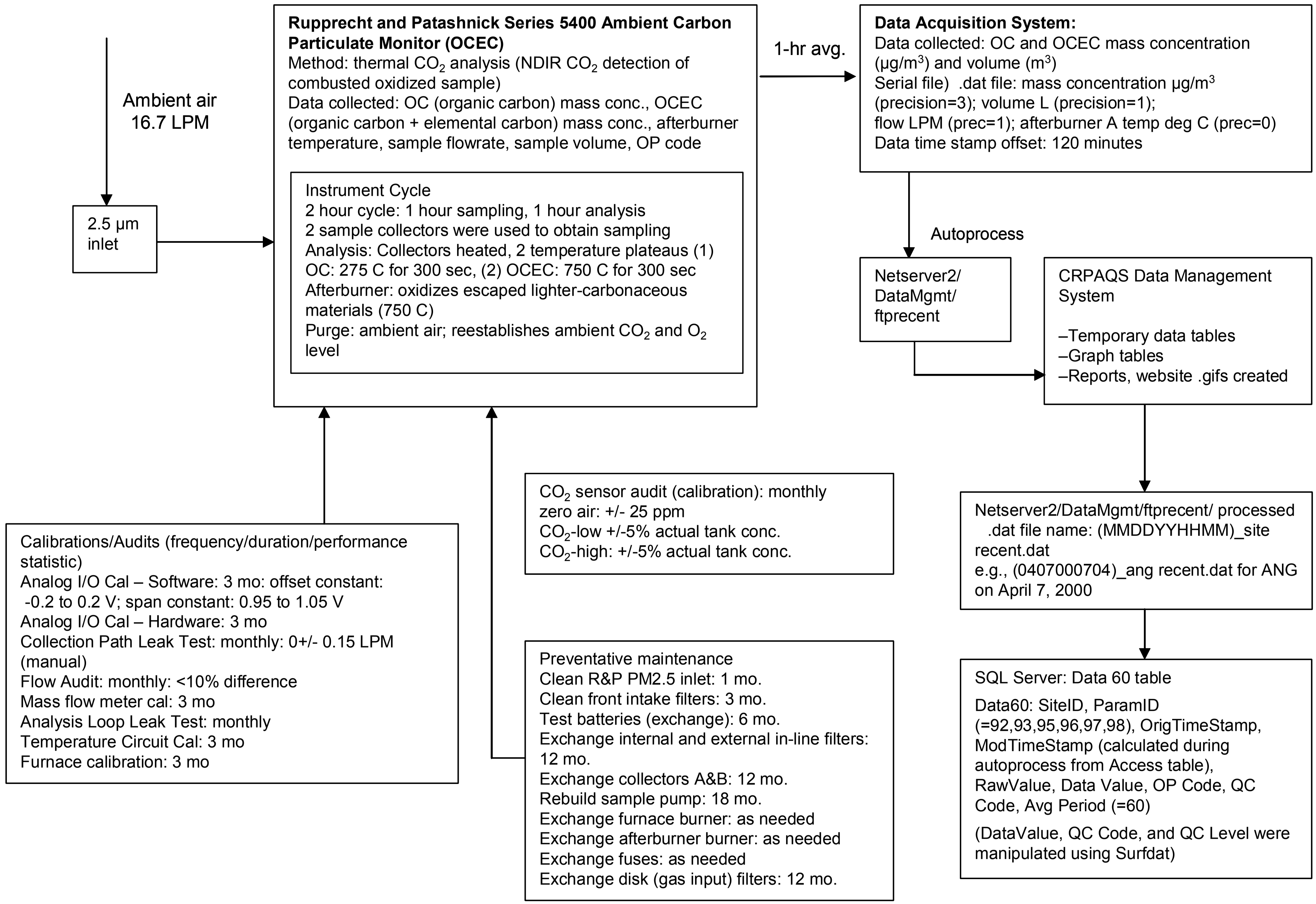


Figure 2-18. Flow chart for the OCEC data.

Table 2-11. Selected R&P 5400 settings. Instrument default and actual CRPAQS site settings are detailed.

	Date	D3 (seconds)	DF (seconds)	P3 (°C)	PF (°C)	AbXT (°C)
Default	---	480	360	340	750	750
BAC	10/04/2000-2/28/2001	300	360	275	750	750
ANGI	2/25/2000- 4/19/2000	600	600	275	750	750
ANGI[a]	4/19/2000- 7/25/2000	600	600	275	750	0
ANGI[a]	7/25/2000-8/10/2000	600	600	275	750	750
ANGI[a]	8/10/2000- 11/03/2000	480	360	275	750	750
ANGI	11/03/2000-2/06/2000	300	360	275	750	750

D3 = OC temperature plateau dwell time
DF = EC temperature plateau dwell time
P3 = OC evolution temperature plateau
PF = EC evolution temperature plateau
AbXT = Afterburner temperature during analysis (X indicates channels A and/or B)

[a] Data were all invalid during this period.

The continuous carbon analyzer draws ambient air for an hour and the ambient air impacts the aerosol on the collection surface. When the collection period (an hour) is finished, the aerosol is combusted. Carbon-containing compounds are oxidized to form CO_2. During the oxidation phase, the system is a closed loop. A non-disperse infrared (LiCor™) sensor downstream of the collector detects CO_2 molecules in the loop; thus, the CO_2 trace should only increase through the analysis period.

In order to differentiate between the CO_2 evolved from OC and the CO_2 evolved from EC and, thus, to determine the concentration of each in the aerosol, the instrument performs the combustion in two steps.

1. The temperature of the collector is raised to an intermediate temperature plateau (denoted P3, third plateau), which lasts for a set amount of time (denoted D3, third dwell time); this step is designed to provide sufficient time and energy for the organic material to combust. Throughout most of the CRPAQS data collection period, the P3 temperature was 275°C, and the dwell time was 300 seconds.
2. The temperature of the collector is elevated to the final plateau (PF) for another set period (DF, final dwell time), designed to combust the remaining aerosol, whose carbon content should be largely elemental. Again, for most of the CRPAQS data collection period, the PF temperature was 750°C and the final dwell time was 360 seconds. The instrument is capable of performing the analysis in up to four steps (hence P3, D3), but this study did not exploit this capability.

Table 2-11 summarizes the instrument default and actual settings by site and date. The effects on the data are discussed in Section 3.5.6.

2.5.7 SO_2

One-minute averaged SO_2 data (ppb) were collected by the Thermo Environmental 43S SO_2 Analyzer at the Bakersfield site. **Figure 2-19** shows the flow of data for this instrument and includes analyzer specifics (averaging period, detection limit, etc.), data acquisition and data storage issues (file names, data table structure), and calibration and maintenance information (nightly, weekly, and multipoint). Data were processed and imported into the 1-minute database on a daily basis during the winter intensive. The SO_2 instrument output analog data only, and these data were transferred to the database via the Recent.DAT file. Nightly calibration information, contained within the CalRecent.DAT file, was plotted alongside the two-day running data in the reports. The span gas concentration changed from 50 to 40 ppb after December 21, 2000, but processing was unaffected. Aside from system-wide issues, there were no problems associated with the processing of the SO_2 parameter.

2.5.8 PAN/NO_2

Data flow for the CE-CERT PAN/NO_2 Gas Chromatograph is displayed in **Figure 2-20**, which details the chromatograph itself (data collected, averaging period); data acquisition (stored precision, file types); data import and storage; calibration schedules (daily, weekly, biweekly, multipoint); and maintenance and other checks. NO_2 (ppb), PAN ($\mu g/m^3$), mode, and baseline information were gathered and deposited into the 1-minute database. The instrument had a built-in automatic calibration system that provided internal zero-span checks, 10 minutes of zero air, and 10 minutes of NO_2 span gas, every two hours. The “mode” parameter provided information about the status of the internal calibrations. The instrument was also attached to the NO_y inlet manifold and received calibration gases from the site calibrator (nightly zero-spans, matrix zeroes, etc.). The PAN/NO_2 analyzer went off-line when the NO_y instrument went off-line. Nightly calibration data were plotted on the same page as the two-day running graphs.

Because of an error in the parameter-naming convention, NO_2 and PAN data for the first few weeks of collection were incorrectly parsed and had to be reprocessed following careful editing of the affected Recent.DAT files. No other major processing problems were encountered.

2.5.9 Nitric Acid

Figure 2-21 describes the data stream for the dual-converter Thermo Environmental 42CY NO_y/HNO_3 Analyzer. Instrument information (data collected, detection limits, averaging period, etc.), calibration (nightly, biweekly, multipoint) schedules, data acquisition issues (stored precision, files involved), data processing and storage details are given. One-minute averaged NO_y and NO_y-HNO_3 data (in ppb) were collected and processed daily and stored in the 1-minute database. This instrument was attached to the NO_y inlet manifold and received calibration gases (zero-span, matrix zero, etc.) from the site calibrator. Nightly zero-span calibration data for the NO_y and NO_y-HNO_3 were plotted adjacent to the two-day running information.

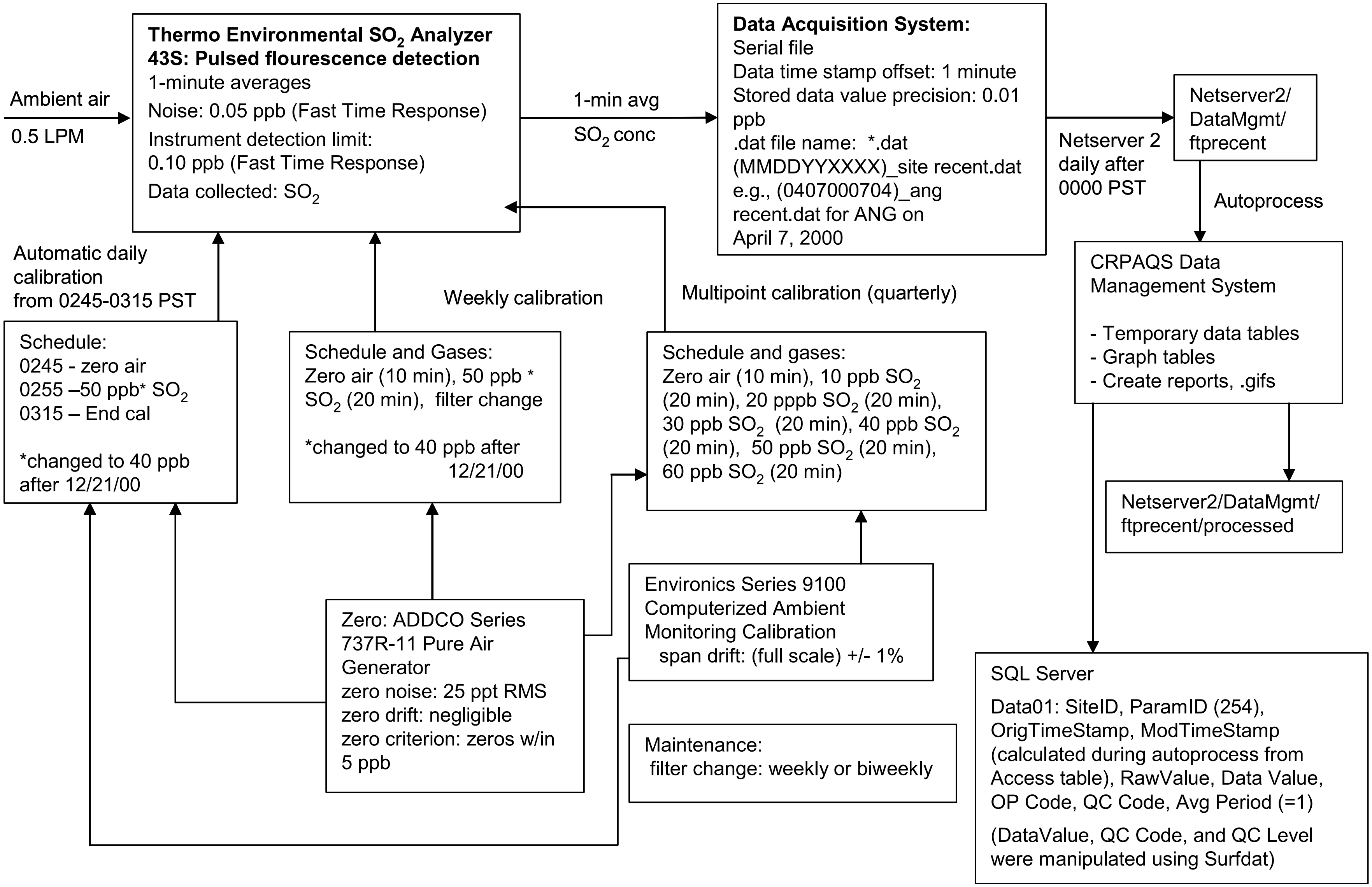


Figure 2-19. Flow chart for SO_2 data.

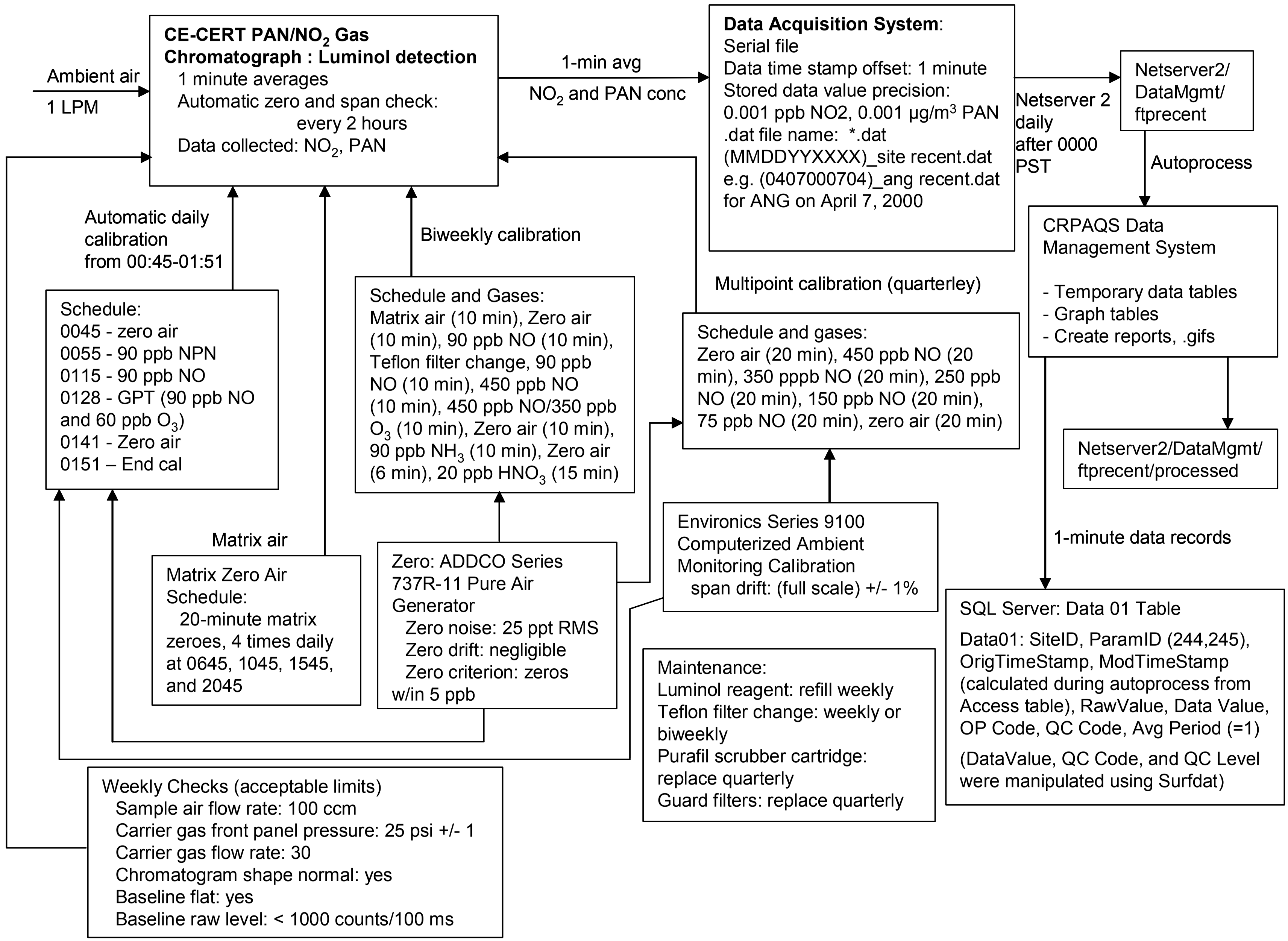


Figure 2-20. Flow chart for PAN/NO_2 data.

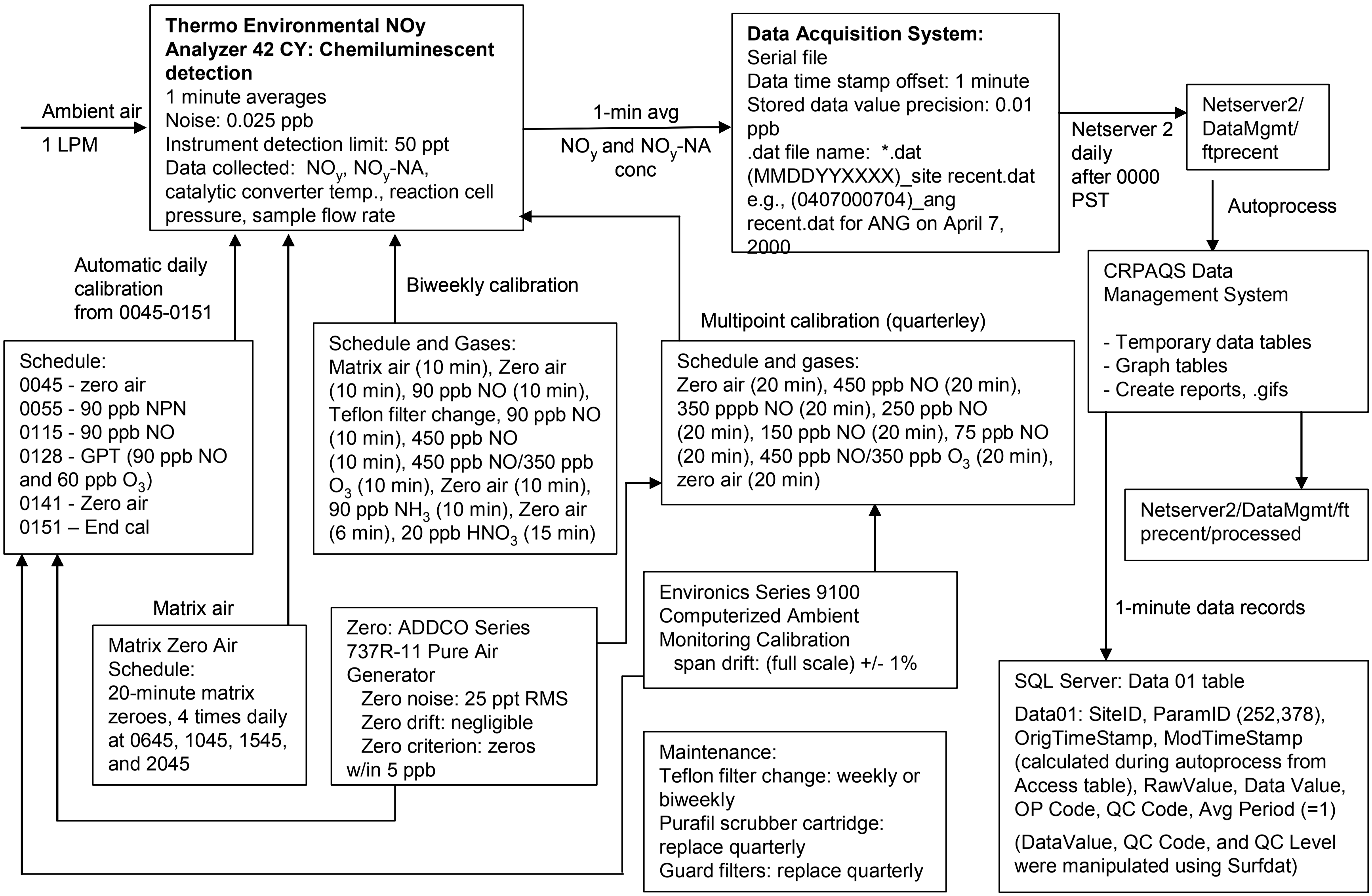


Figure 2-21. Flow chart for nitric acid (NA) data.

At Angiola, a naming convention error was discovered in the first few days of instrument operation after the data parser failed to recognize the NO_y-HNO_3 parameter. The affected Recent.DAT files were edited, and NO_y-HNO_3 data were reprocessed without a problem. However, a similar problem occurred at the DAS level at the Sierra Nevada Foothills site when the same values (corresponding to NO_y-HNO_3 data) were pulled in for both parameters, resulting in brief data loss.

Prior to data validation, a new parameter was created in the 1-minute database corresponding to calculated nitric acid. For each minute in which both a NO_y and NO_y-HNO_3 data record were available, a difference was taken and a new nitric acid record produced.

2.5.10 Continuous Nitrate

Details of the data flow for the Rupprecht and Patashnick 8400N Ambient Particulate Nitrate Monitor are found in **Figure 2-22**. A number of parameters were collected and imported into the 10-minute database on a daily basis, but only NO_3 ($\mu g/m^3$) and NO_x (ppb) information were plotted. The pressure data from the collection/vaporization cell (Rcell parameter) was not initially imported; however, because of its importance, the measurement expert issued instructions to reprocess and add the data at a later date. With this newly commercialized instrument, numerous missing data gaps reflect the myriad problems encountered.

2.5.11 Continuous Sulfate

The Rupprecht and Patashnick 8400S Ambient Particulate Sulfate Monitor was not available for use until the final weeks of the winter intensive and was plagued with similar operational problems experienced with the nitrate instrument. **Figure 2-23** illustrates the flow of data from sampling to data storage for the continuous sulfate sampler. A number of parameters were processed and imported into the 10-minute database, but only the SO_2 (ppb) and SO_4 ($\mu g/m^3$) data were plotted. Data processing for this instrument was relatively smooth, but there are a number of data gaps in the database resulting from periods when the instrument failed.

2.5.12 Particle Counts by SMPS

The size distribution of fine and ultrafine aerosol (10 to 400 nm, or 0.01 to 0.4 μm) was measured by a TSI Scanning Mobility Particle Sizer (SMPS), located in the trailer at the Angiola site and configured as part of the Particle Sizing System, which also included a Climet OPC and a PMS Lasair OPC. The SMPS size distribution was reported as particle concentrations in 51 size bins. Particle concentrations, sample flow rate, sheath flow rate, sheath stability, differential mobility analyzer (DMA) absolute pressure, and instrument date and time were reported daily in the 5-minute database. **Table 2-12** details the size-cuts of the 51 bins.

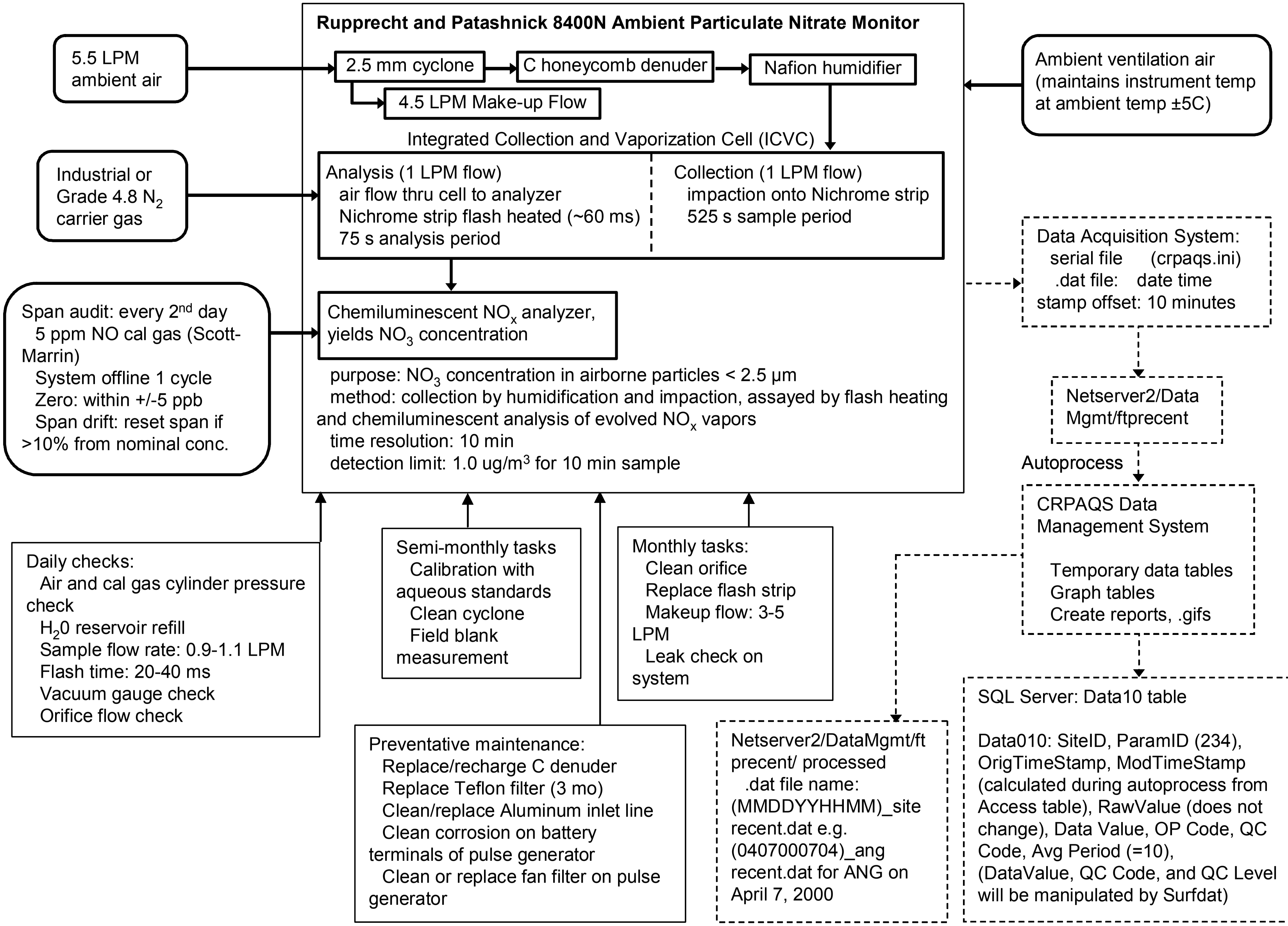


Figure 2-22. Flow diagram for the continuous nitrate monitor (prepared by B. Kirby, ADI).

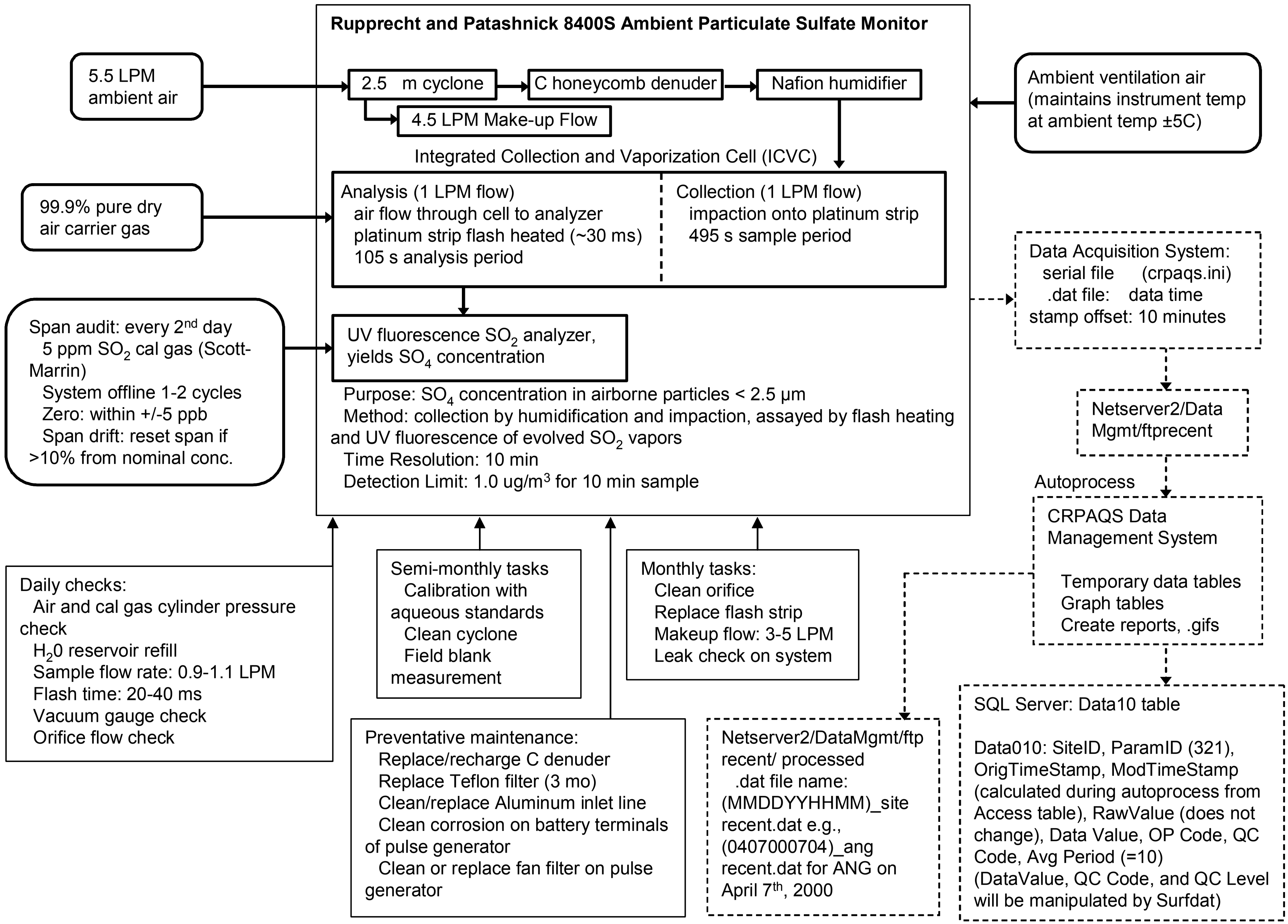


Figure 2-23. Flow diagram for the continuous sulfate monitor (prepared by B. Kirby, ADI).

Table 2-12. SMPS S/N 8083 bin sizes. D_P indicates particle diameter in nm.

Channel	D_P low	D_P high
1	10.0	10.7
2	10.7	11.5
3	11.5	12.4
4	12.4	13.3
5	13.3	14.3
6	14.3	15.4
7	15.4	16.5
8	16.5	17.8
9	17.8	19.1
10	19.1	20.5
11	20.5	22.1
12	22.1	23.7
13	23.7	25.5
14	25.5	27.4
15	27.4	29.4
16	29.4	31.6
17	31.6	34.0
18	34.0	36.5
19	36.5	39.2
20	39.2	42.2
21	42.2	45.3
22	45.3	48.7
23	48.7	52.3
24	52.3	56.2
25	56.2	60.4
26	60.4	64.9
27	64.9	69.8

Channel	D_P low	D_P high
28	69.8	75.0
29	75.0	80.6
30	80.6	86.6
31	86.6	93.1
32	93.1	100
33	100	107
34	107	115
35	115	124
36	124	133
37	133	143
38	143	154
39	154	165
40	165	178
41	178	191
42	191	205
43	205	221
44	221	237
45	237	255
46	255	274
47	274	294
48	294	316
49	316	340
50	340	365
51	365	392

The TSI SMPS Model 3936L10 consists of an Electrostatic Classifier (ESC) Model 3080 coupled to a Long Differential Mobility Analyzer (DMA) Model 3081 followed by a Condensation Particle Counter (CPC) Model 3010. The system is controlled and data recorded by software installed on a Windows 95 computer.

The 1 L/min sample stream first passes through an impactor (0.0508 cm orifice) on the front of the ESC to remove particles that are larger than 410 nm. The particles then pass through a radioactive charger/neutralizer (Model 3077, Kr-85, 2 mCi) inside the ESC, where they interact with bipolar ions to acquire a known steady-state charge distribution. Due to random collisions, some of the entering particles acquire a positive charge, others a negative charge, and others remain neutral. In addition, a small fraction of the larger particles acquire multiple charges.

The 1 L/min sample flow then passes into the DMA, entering along the inner wall at the top of a vertical cylinder where it smoothly joins a co-current 7 L/min particle-free sheath flow between it and a concentric center rod. As these two streams flow laminarly down the length of the annular gap between the outer cylinder and the inner rod, positively charged particles are pulled across the clean sheath flow toward the negatively charged center rod. At the same time, while negatively charged particles precipitate on the grounded inner cylinder wall and neutral particles remain radially undeflected. The positively charged particles are effectively sorted by electrical mobility along the center rod.

Particles within a narrow range of electrical mobility are drawn out of the main flow in a 1 L/min flow through a narrow circumferential slit near the bottom of the center rod. Most of these selected particles have one positive charge with a relatively small fraction having two (or more) positive charges. The CPC then measures the concentration of the selected aerosol by condensing butanol vapor onto the particles and growing them to a size large enough to be detected and counted optically as they pass through a laser beam. Downstream of the optics is a critical orifice and a dedicated pump which maintain the sample flow through the CPC at 1 L/min.

Over a period of 135 seconds the center rod voltage and corresponding selected particle diameter is scanned from the minimum to the maximum (followed by a 15 second down scan) and the CPC count is recorded in tenth-of-a-second increments. Theoretical relationships are used to convert from scan time and voltage, to electric mobility, to particle diameter. Knowledge of the charge distribution is used to convert measured concentrations of charged particles to total concentration at each particle size. Correction for multiply-charged particles is also possible. Within each 5-minute period, two scans are made and summed and the size distribution reported with resolution of 32 channels per decade of particle diameter.

Overall control of the SMPS system is handled by the TSI SMPS program through serial communication with the CPC. Scan parameters (DMA voltage limits and scan times) are sent to the CPC which in turn sends an analog exponential control voltage ramp to the ESC which amplifies the signal by a factor of 1000 and applies it to the DMA center rod. The ESC also controls the DMA recirculating sheath flow and measures the temperature and absolute pressure of that flow as well as the pressure drop across the impactor. The sealed recirculation loop for the sheath flow ensures that the aerosol flow into the DMA is equal to the 1 L/min outflow fixed by the CPC critical orifice. During the voltage scan the CPC automatically sends tenth-second particle counts back to the SMPS program, which stores the raw counts in one file per 5-minute sample and calculates the size distribution as described above. However, the TSI SMPS program does not communicate with the ESC but uses fixed standard operating parameters for these calculations. The TSI SMPS program also is not able to transmit the size distribution serially to the site DAS.

As a result, a companion program (SMPSDAT) was written by ADI (in Microsoft QuickBASIC). SMPSDAT communicates serially with both the ESC and the site DAS and both programs ran concurrently on the SMPS system computer. At startup the SMPSDAT program sends setup parameters (e.g., sheath flow set point) to the ESC then waits for raw sample data files from the SMPS program. When it finds a new data file, SMPSDAT: reads and records operating parameters (temperatures, pressures, flows and flags) from the ESC; reads the data file and calculates the size distribution using current operating conditions; sends the processed data to the DAS and to a local monthly ASCII file; and moves, renames and zips the raw data file into a monthly ZIP file.

Not all of the shortcomings of the TSI SMPS software could be overcome. Once started, the SMPS program could only take 999 consecutive 5-minute samples (~3.5 days) before an operator would need to restart the data collection or the TSI SMPS program would stop. Thus, the SMPS system required operator intervention at least every third day or data were lost. It was not possible to trigger the SMPS sample start with a serial command from the site DAS. Instead, the SMPS ran according to the SMPS computer clock independent of the DAS clock. Due to program delays, the period from sample start to sample start was a fraction of a second more than 5 minutes causing the sample start times to "creep" with respect to the SMPS computer clock. The SMPS computer clock in turn "crept" with respect to the DAS clock. To compensate for the accumulating time creep, the site operator manually resynchronized the SMPS computer clock to the DAS clock, and the SMPS program to the SMPS computer clock, when restarting the data collection every third day.

Because the parameters could not all be plotted at the same time without seriously impairing graph clarity, only particle count data from the 0.05 μm bin and the associated sample flow were viewed by the STI data technician during the daily quality control checks. Data from the smallest size bin had been graphed previous to August 17, 2000, but abandoned due to low particle counts in that bin.

Data at the Angiola site for the largest size sample bin were reprocessed for time periods between December 24, 2000 and January 26, 2001, when it was discovered that the data were missing. There were no other major problems with the initial processing for this instrument. Data were post-processed by ADI. STI converted these data from ADI's delivered format and brought them into STI's database for final processing.

Data flow for the SMPS is shown in **Figure 2-24**.

2.5.13 Particle Counts by Climet OPC

Coarse particle (0.3 to 10 μm) sizing was performed by three Climet Instruments CI-500 Spectro 0.3 Optical Particle Counters (OPC) located at ground level, 50 m agl, and 100 m agl at the Angiola site. The size distribution was reported as particle counts in 16 bins, and this count information, along with sample flow, was imported into the 5-minute database on a daily basis. Each Climet OPC instrument is unique in its bins' size-cuts. **Table 2-13** details information about the individual instruments' bins.

The size cuts are relatively independent of flow rate. Although flow rate affects sample volume and, consequently, particle counts, the size cuts themselves remain constant. However, the counts in the Climet's larger bins are strongly affected by flow rate because the flow rate determines the size selectivity of the inlet—in this case, a PM_{10} sharp-cut cyclone. When the flow through the inlet decreases significantly, larger particles enter the system. When the flow through the inlet increases significantly, the size cut of the inlet is decreased.

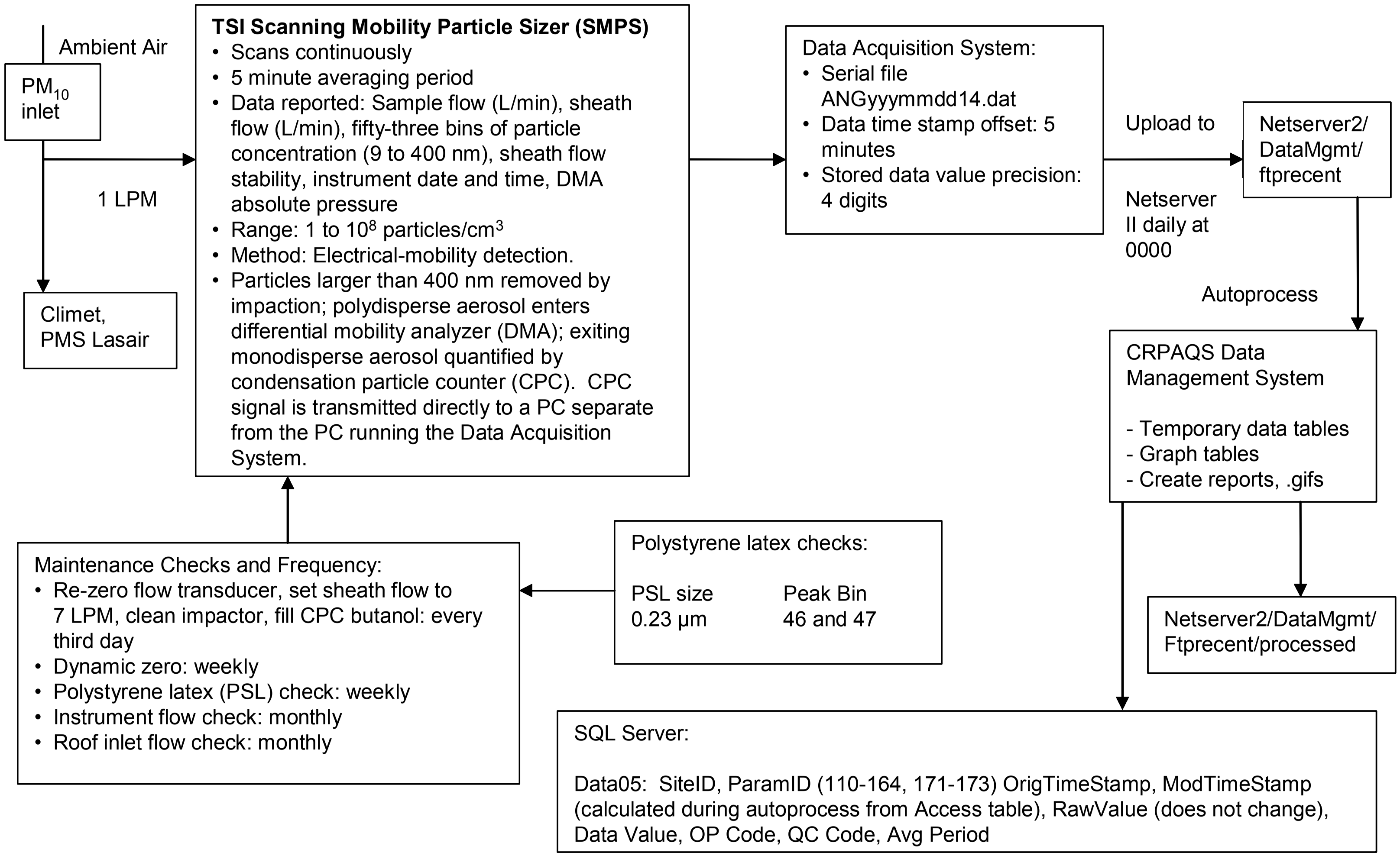


Figure 2-24. Flow chart for SMPS particle count data.

Table 2-13. Calibration-determined channel cut points for Climet OPCs. D_P = particle diameter in micrometers (µm).

Channel	Climet S/N 978182		Climet S/N 990246		Climet S/N 990247	
	D_P, Low	D_P, High	D_P, Low	D_P, High	D_P, Low	D_P, High
1	0.323	0.410	0.362	0.449	0.360	0.447
2	0.410	0.521	0.449	0.556	0.447	0.556
3	0.521	0.661	0.556	0.688	0.556	0.691
4	0.661	0.839	0.688	0.852	0.691	0.858
5	0.839	1.06	0.852	1.06	0.858	1.07
6	1.06	1.35	1.06	1.31	1.07	1.32
7	1.35	1.72	1.31	1.62	1.32	1.65
8	1.72	2.18	1.62	2.01	1.65	2.05
9	2.18	2.76	2.01	2.49	2.05	2.54
10	2.76	3.51	2.49	3.08	2.54	3.16
11	3.51	4.45	3.08	3.81	3.16	3.92
12	4.45	5.65	3.81	4.72	3.92	4.87
13	5.65	7.18	4.72	5.85	4.87	6.06
14	7.18	9.11	5.85	7.24	6.06	7.52
15	9.11	11.56	7.24	8.97	7.52	9.35
16	11.56		8.97	(10)	9.35	

Although the instruments were at different heights, data from the three Climet OPCs were included in the same daily Recent.DAT file. The parameters of each instrument were distinguished from one another by proper naming and referencing in the database. Count data only from Channel 1 (smallest size-cut) and sample flow rate were plotted and viewed during the daily QC check.

The Angiola Tower instruments were configured differently from the instrument in the trailer, the effect of which was investigated in intercomparison testing. In its ground configuration, the Climet OPC was part of the particle sizing system, which included the TSI SMPS and PMS Lasair OPC. Thus, it pulled its sample flow from an air stream that also fed the SMPS and Lasair OPC. For this reason, among others, Climet OPC instruments at the various levels were often compared with one another and examined in combination with nephelometer data from the same height during the daily QC checks.

Data flow for the Climet OPC is shown in **Figure 2-25**.

2.5.14 Particle Count by PMS Lasair OPC

Mid-range particle sizing (0.1 to 2 µm) was performed by the PMS Lasair OPC at ground level at the Angiola site. This instrument provided a particle distribution via count information for 8 size-selective bins. **Table 2-14** lists specific bin size cuts. Although count data for the 8 bins were imported daily into the 5-minute database, only data from Bin 1 (smallest size cut) and the sample volume (L) were plotted and reviewed for the daily QC checks.

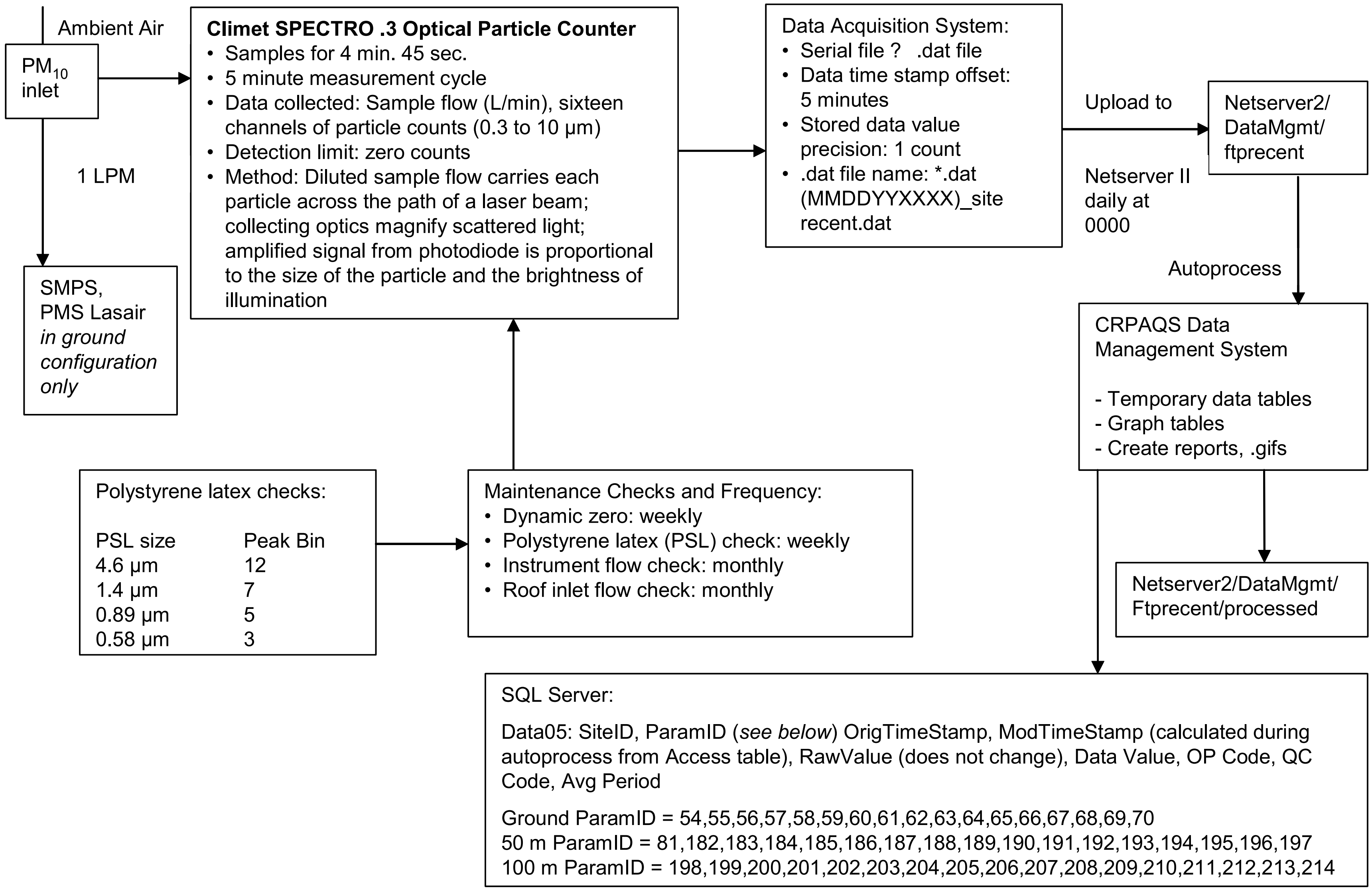


Figure 2-25. Flow chart for Climet OPC particle count data.

Table 2-14. Calibration-determined channel cut points for PMS Lasair OPC S/N 12573. D_P = particle diameter in micrometers (µm).

Channel	D_P, Low	D_P, High
1	0.111	0.221
2	0.221	0.316
3	0.316	0.431
4	0.431	0.600
5	0.600	0.843
6	0.843	1.18
7	1.18	2.02
8	2.02	

The size-cuts are relatively independent of sample flow rate but extremely dependent on the ratio of sheath flow to sample flow.

Data flow for the PMS Lasair OPC is shown in **Figure 2-26**.

2.6 ARCHIVING

Part of preparing a traceable database is to properly keep track of supporting documents. Documentation included daily site log forms, equipment logbooks, DAS and instrument files, instrument task sheets, operating manuals, SOPs, and equipment lists. All these items are archived at STI in addition to being available in the field. **Table 2-15** lists the deliverable products and other archived items.

2.6.1 Copies of Daily Site Log Forms

The site operators used site log forms as checklists when they arrived at and left a site. The forms provided an overview of the operators' actions and observations for the day and were designed to be useful in the daily data QC checks.

STI received from the field daily faxes of the site log form containing operator notes recorded at each site. These forms were retrieved from the fax machine, three-hole punched, and placed in a site binder. At the end of the week, the STI data management technician verified that all the site log forms were obtained (Monday through Friday). The technician notified the field manager if there were any missing site log forms.

2.6.2 Copies of Equipment Logbooks

Separate logbooks were used for each piece of equipment at each site. These logbooks document any actions or observations regarding a piece of equipment including maintenance, troubleshooting, calibrations, etc.

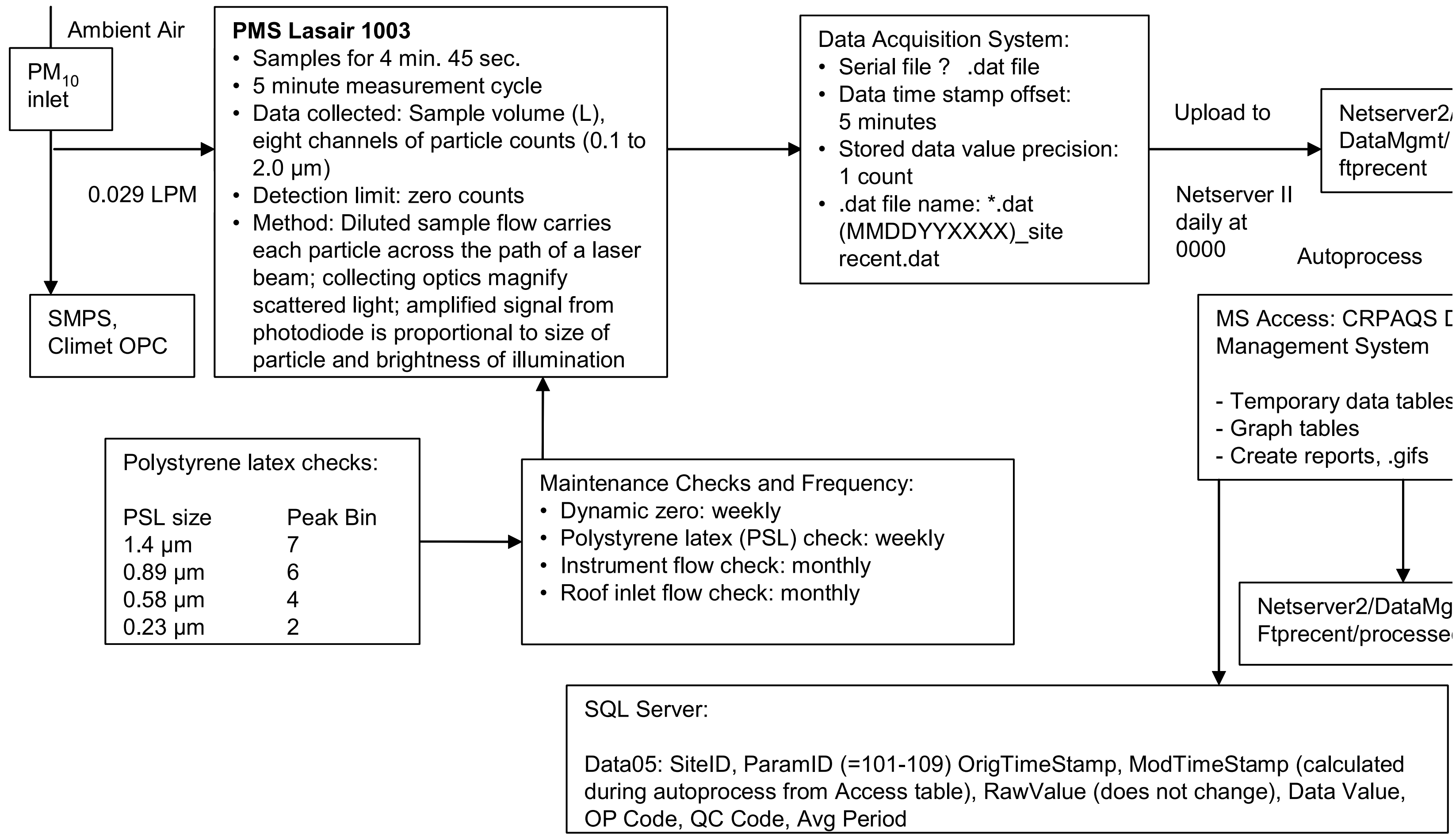


Figure 2-26. Flow chart for the PMS Lasair OPC particle count data.

Table 2-15. STI data management products.

Data or Information	Description	Location
Recent.DAT files	DAS daily output organized by date	Zipped files CDs or DVDs created and stored at STI
Instrument hard drive files	Copies of data from hard drives of selected instruments	Various floppy disks copied to CD and stored at STI
Daily raw data graphs	Two hard copy versions existed during the study (at the STI and Bakersfield field offices) *.GIF files	Hard copies filed at STI *.GIF files on server backed up onto CD
MS Access database	Front end to the database containing all tables and processes	Backed up on CD Resides on more than one computer at STI
SQL server database	Entire database of all data retrieved and stored at STI	On server Backed up onto CDs or DVDs Copy delivered to ARB
Raw CDF files	Export of raw data to SurfDat	Zipped on data technician computer Copied to CDs and stored at STI
Site systems audit data		Resides on STI network; routinely backed up
Performance audit data		Resides on STI network; routinely backed up
Modified CDF files and SurfDat logs	After data validation, files ready for importing back to database	Zipped on data technician computer Backed up on CDs
SurfDat	Version of software used to validate data	Backed up on CDs
Export routine	MS Access program used to prepare deliverable to ARB	Backed up on CDs
Deliverable data	ARB formatted data	Backed up on CDs Delivered to ARB
Site and instrument logbooks	Operator notes	Original and complete copies filed at STI
Strip charts	Instrument hard copy output	Original and complete copies filed at STI
Control charts	Daily zero, span, and maximum value data for each instrument	Original files at STI
Calibration and off-line summaries	Prepared during data validation	Original notes filed at STI
Transfer standard certification information		Original notes filed at STI

On Mondays during the field study, STI received copies of the instrument logbooks from the Angiola, Bakersfield, Sierra Nevada Foothills, and Bethel Island sites by overnight delivery or fax. The data management technician filed the log entries by instrument and site in a dedicated filing cabinet. Logbook copies were kept in chronological order. Whenever new pages were received, the technician checked for continuity. If missing pages were identified, the field manager was notified

2.6.3 Instrument Initialization and Calibration Files

All of the programmable features of the data acquisition program (DAP) except calibration command files were configured using the initialization file (CRPAQS.INI). This self-documenting file was read by the DAP once upon starting, and the DAP immediately began collecting data according to the instructions in the file. Changes to the initialization files were made only by authorization of the field manager.

The initialization file (*.INI) had five main section types:

1. "Configuration" set the site ID, defined the DAS hardware, identified the calibration file directory, and defined the truncation size of the Recent.DAT file.
2. "AutoCal" set the execution start time and specified the *.CAL files used in regularly scheduled calibrations.
3. "Instruments" identified all the instruments connected to the DAS.
4. "Serial Template" was used as a template to define serial communication procedures with an instrument identified in the "Instruments" section.
5. "Analog Template" was used as a template to configure analog input channels from an instrument identified in the "Instruments" section.

During the field study, changes to the initialization files were coordinated with the field manager and the data manager. Current and old initialization files were archived. The version of the initialization file and the latest changes, in reverse chronological order, were included in notes at the beginning of the file.

The initialization file referenced *.CAL script information for the automated calibrations that were to occur on a scheduled basis. For all other calibrations, the site operator would manually initiate the appropriate *.CAL files by selecting it from the DAS screen. Current and old *.CAL files were archived, and changes to subsequent *.CAL files were noted at the top of the file.

2.6.4 Instrument Operating Manuals

STI maintained a separate copy of each instrument operating manual in the field and at its office in Petaluma.

2.6.5 Instrument Standard Operating Procedures

Standard operating procedures (SOPs) for each instrument were maintained electronically at the STI office and available in Wittig et al. (2003); printed copies were available at the field offices.

2.7 DATABASE MAINTENANCE

Database backups to tape were created weekly during the field study and data validation. Once the entire data set was subjected to Level 1 validation and exported to ARB, we created CD or DVD backup copies for storage at STI and delivery to ARB.

3. DATA VALIDATION

3.1 OVERVIEW

Quality assurance (QA) is a CRPAQS project management responsibility that integrates QC, quality auditing, and measurement method validation into the measurement process. QC is intended to prevent, identify, correct, and define measurement difficulties and to provide the QC test data needed to quantify the precision, accuracy, and validity of the data. QC activities include (1) preparing SOPs and other project documentation to be followed during sampling, analysis, and data processing; (2) testing equipment for acceptance, repairing equipment as needed, and maintaining spare parts; (3) providing operator training, supervision, and support; and (4) performing periodic calibrations and performance tests which include blank and replicate analyses.

QC activities in the field are described by Wittig et al. (2003). QC activities performed as a part of the daily data management included

- Printing and reviewing the previous days' data and instrument response checks by the data technician. Outliers were identified and reviewed to ensure the data were physically realistic.
- Maintaining redundancy of data recording using log sheets, data printouts, and multiple magnetic recording media.

The daily data management QC activities were discussed in Section 2. This section contains a discussion of STI data validation activities and technical approach prior to data delivery.

3.2 DATA VALIDATION LEVELS

"The purpose of data validation is to detect and then verify any data values that may not represent actual air quality conditions at the sampling station" (U.S. Environmental Protection Agency, 1980). Thorough data validation is critical because serious errors in data analysis and modeling results can be caused by erroneous individual data values. In order to minimize unnecessary errors and uncertainties, all values in the data set were reviewed, evaluated, and flagged. The identification of outliers, errors, or biases is typically carried out in several stages or validation levels as defined by Watson et al. (1998).

- **Level 0:** These data are obtained directly from the data loggers that acquire data in the field. Averaging times represent the minimum intervals recorded by the data logger, which do not necessarily correspond to the averaging periods specified for the database files. Level 0 data have not been edited for instrument downtime, nor have procedural adjustments for baseline and span changes been applied. Level 0 data are not contained in the database delivered to ARB.
- **Level 1A:** These data have passed several validation tests applied by the measurement expert prior to data submission. The general features of Level 1A are (1) flagging and removal of data values and replacement with -999 when monitoring instruments did not function within procedural tolerances; (2) flagging measurements when significant deviations from measurement assumptions have occurred; (3) verifying computer file

entries against data sheets; (4) replacing data from a backup DAS in the event of failure of the primary system; (5) adjusting measurement values for quantifiable baseline and span or interference biases; and (6) identifying, investigating, and flagging data that are beyond reasonable bounds or that are unrepresentative of the variable being measured (e.g., high light scattering associated with adverse weather).

- **Level 1B:** These data have met consistency tests to verify that file naming conventions, data formats, site codes, variable names, reporting units, validation flags, and missing value codes are consistent with project conventions. Discrepancies are reported to the measurement expert for remediation. When the received files are consistent, reasonability tests are applied that include identifying data values that (1) are outside of a specified minimum or maximum value; (2) change by more than a specified amount from one sample to the next; and (3) do not change over a specified period. Data identified by these filters are individually examined and verified with the data supplier. Obvious outliers (e.g., high solar radiation at midnight, 300°C temperature) are invalidated. Others may be invalidated or flagged based on the results of the investigation. The bounds used in these tests are determined in cooperation with measurement experts and network operators.

STI prepared the data to Level 1B for delivery to the ARB CRPAQS Data Manager. Every effort was expended to retain as much data and information in the database as possible. Invalidated data were assigned a null code to replace the "data" value when the data were not representative of ambient conditions, for example, calibration data and instrument maintenance problems. Data gaps (e.g., when instruments were off-line) were filled with missing data codes. An electronic log was kept of all actions taken on the individual values in the data set.

Dr. Watson provided additional data validation guidance for Level 1:

- Remove clearly invalid data. Common reasons for invalidating measurements are (1) instrument not operating; (2) instrument operating, but not acquiring valid data (as indicated by performance tests, broken parts, excessive leaks, recorded values beyond instrument output limits [electronic noise]); or (3) <75% valid data for the sample duration (<4 minutes for 5-minute averages, <45 minutes for 1-hr averages, <16 hours for 24-hr averages).
- Do not discard negative or very large values, unless they correspond to indicators of invalid data as noted above. These may be real readings for short-duration (five minute) data.
- Apply adjustments where calibration offsets or span biases are known to be consistent for a specific time period.
- Apply flags for zero deviations beyond a specified amount.
- Apply flags for span deviations beyond 15%.
- Estimate precisions for the most basic averaging period from zero drift and span reproducibility. The data base system should calculate standard errors when calculating hourly averages from 5-min averages and 24-hour averages from hourly averages.
- Apply flags for deviations from procedures, sample contamination, and unusual events.

STI followed these guidelines in the data validation steps.

3.3 DATA VALIDATION PROCESS

3.3.1 Overview

The data validation process began with the site operators. Each day that operators visited the sites, they reviewed graphs of recent data. This review augmented their normal operational checks performed according to a checklist and the instrument SOPs. The site operators' vigilance and adherence to SOPs were critical in obtaining good quality data in the field. Once the data were delivered to STI, the data management technician followed several data validation steps:

- Reviewed site and instrument logs from the field staff to assess and summarize instrument off-line times, calibrations, maintenance, etc.
- Reviewed flow audit information.
- Queried the database to obtain all calibration, zero, and span information recorded in the DAS.
- Prepared annotated plots of zero, span, and calibration information.
- Performed general troubleshooting.

The calibration information and plots were then provided to the measurement experts for review. The measurement experts recommended adjustments to the raw data based on their review.

Two important techniques were used to ensure high-quality data validation. First, measurement experts (see Table 1-2) worked with the data management team to develop screening criteria used to identify potentially suspect or invalid data. These criteria were manually applied. The second approach to data validation was the use of in-house data visualization software developed to facilitate graphical review of all data points. The software allows the reviewer to plot several parameters at a time (e.g., ozone, NO, and NO_y; BC reported by all seven wavelengths of the Aethalometer); change QC flags; and apply zero offsets or calibration slope adjustments. All actions were recorded in a log file with entries that uniquely identified the altered data records and documented what changes were made.

Flow charts shown in **Figures 3-1 through 3-6** illustrate many of the steps taken to validate the data. Figure 3-1 shows the validation process for the 1-minute gaseous data. Figure 3-2 shows the data validation approach for 1-minute and 5-minute nephelometer data. Five-minute Aethalometer validation is shown in Figure 3-3 while the 60-minute averaged particulate data validation is summarized in Figure 3-4. Figures 3-5 and 3-6 show validation procedures for the nitrate and sulfate data. Validation procedures for the other instruments are discussed in the appropriate sections.

Data validation for continuous monitors, for example, was conducted in a series of steps:

- Graphically reviewing time series of pollutant concentrations paying particular attention to times before, during, and after calibrations, maintenance, and other off-line periods. Inspect data spikes, dips, and outliers. Plotting complementary data together (e.g., ozone and NO/NO_y).

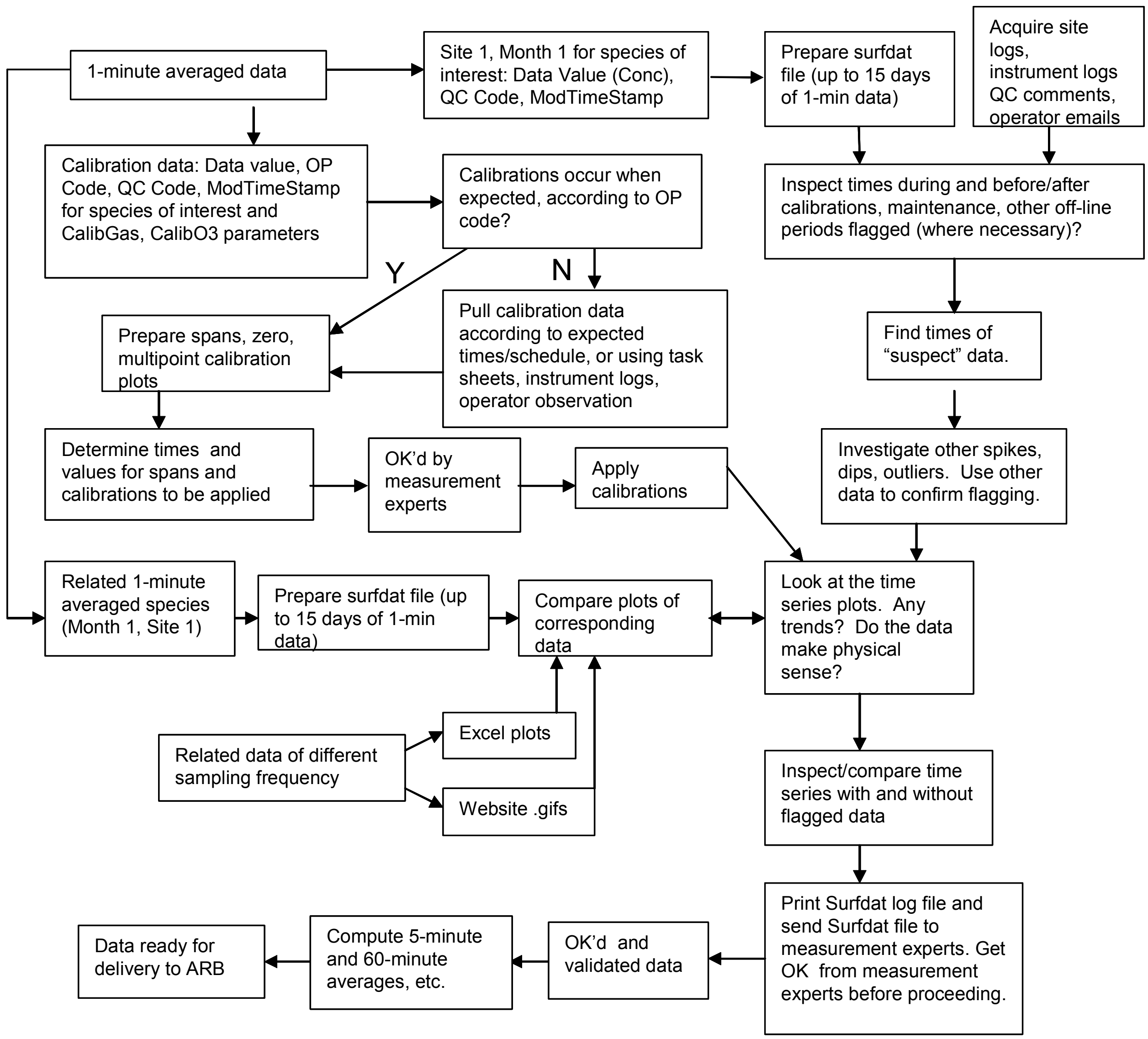


Figure 3-1. Data validation flow chart for 1-minute gaseous measurements (ozone, NO/NO_y, SO_2, PAN/NO_2 and nitric acid).

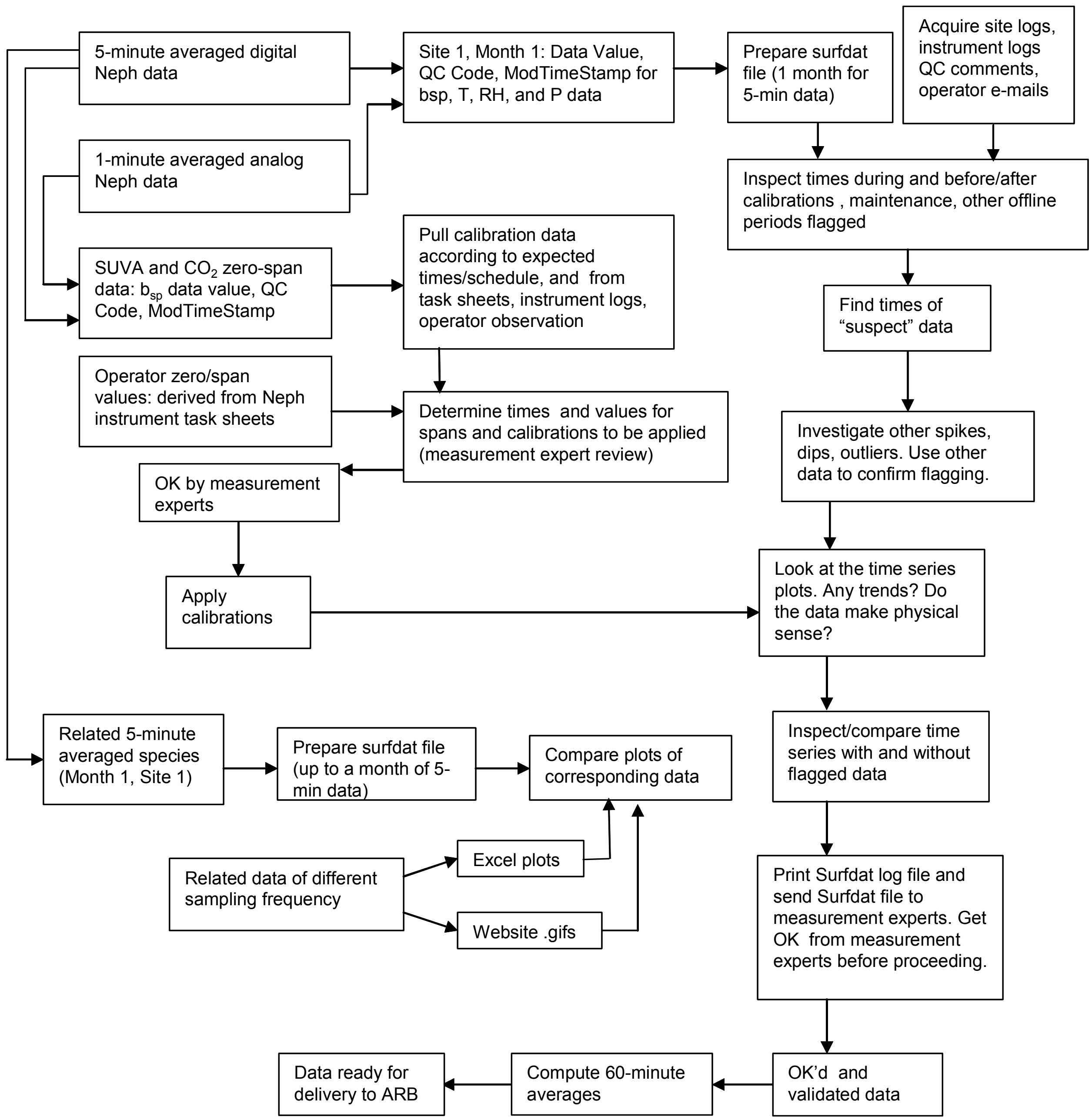


Figure 3-2. Data validation flow chart for 5-minute nephelometer data.

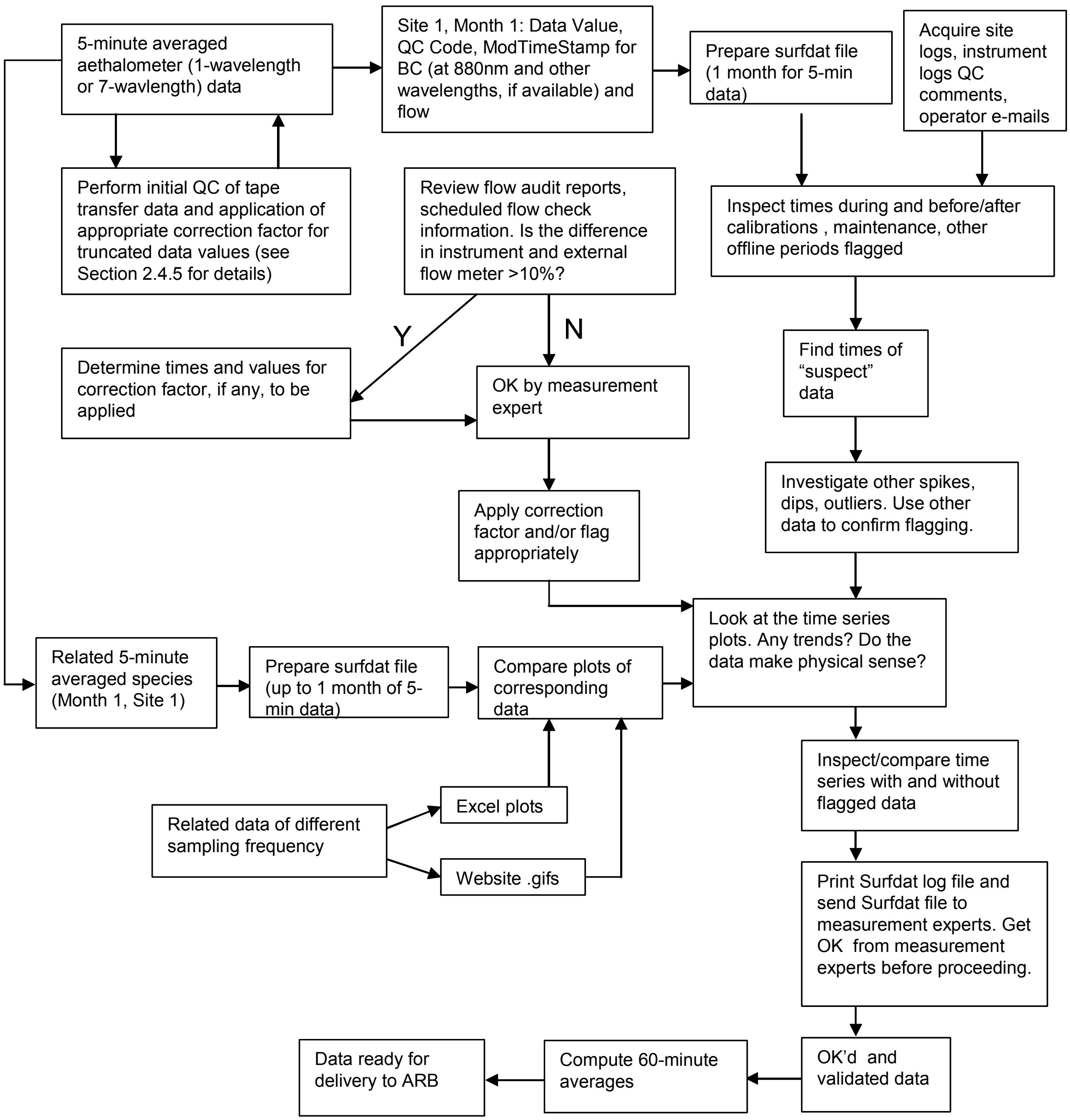


Figure 3-3. Data validation flow chart for 5-minute Aethalometer data.

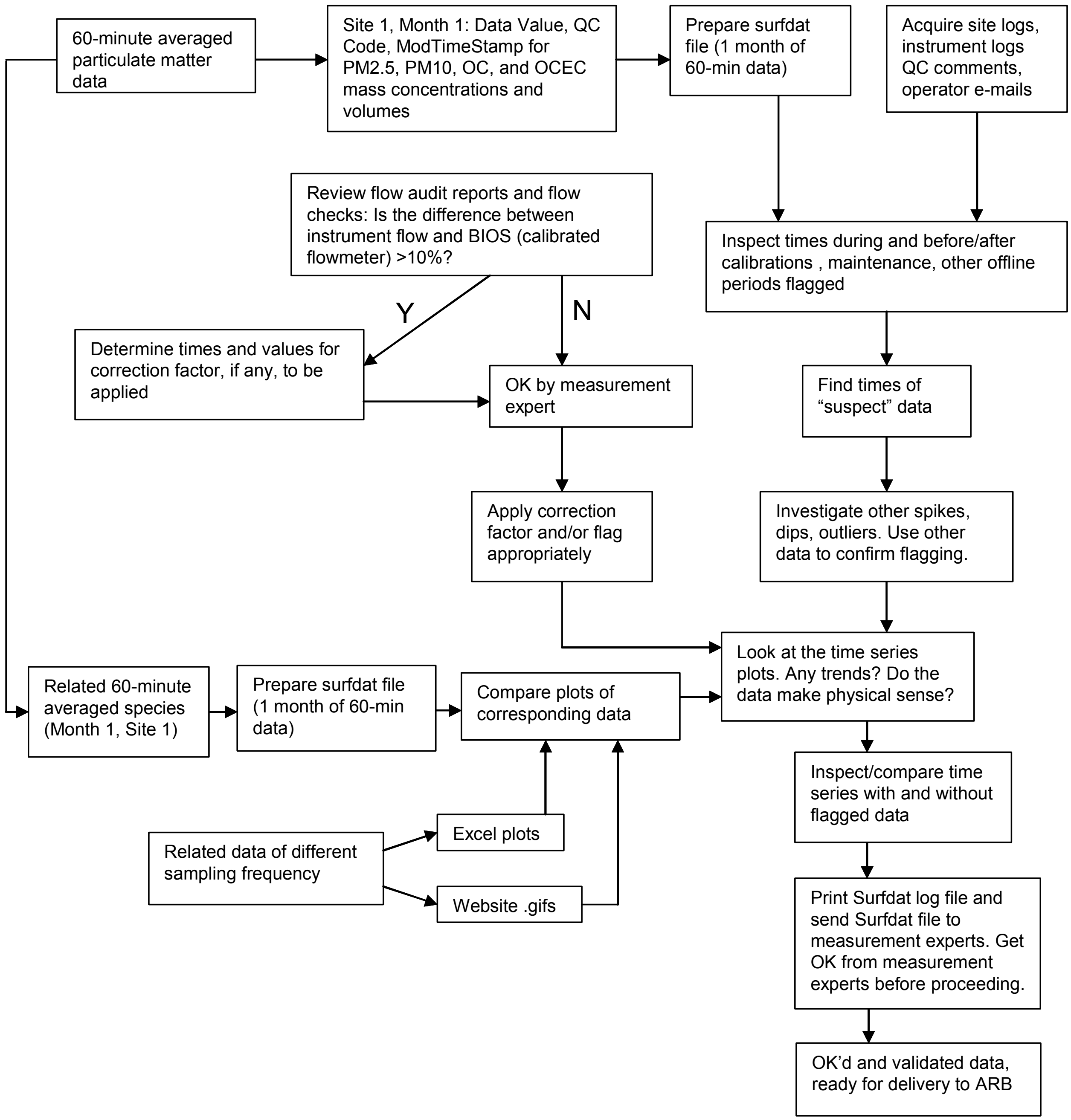


Figure 3-4. Data validation flow chart for 60-minute particulate matter data.

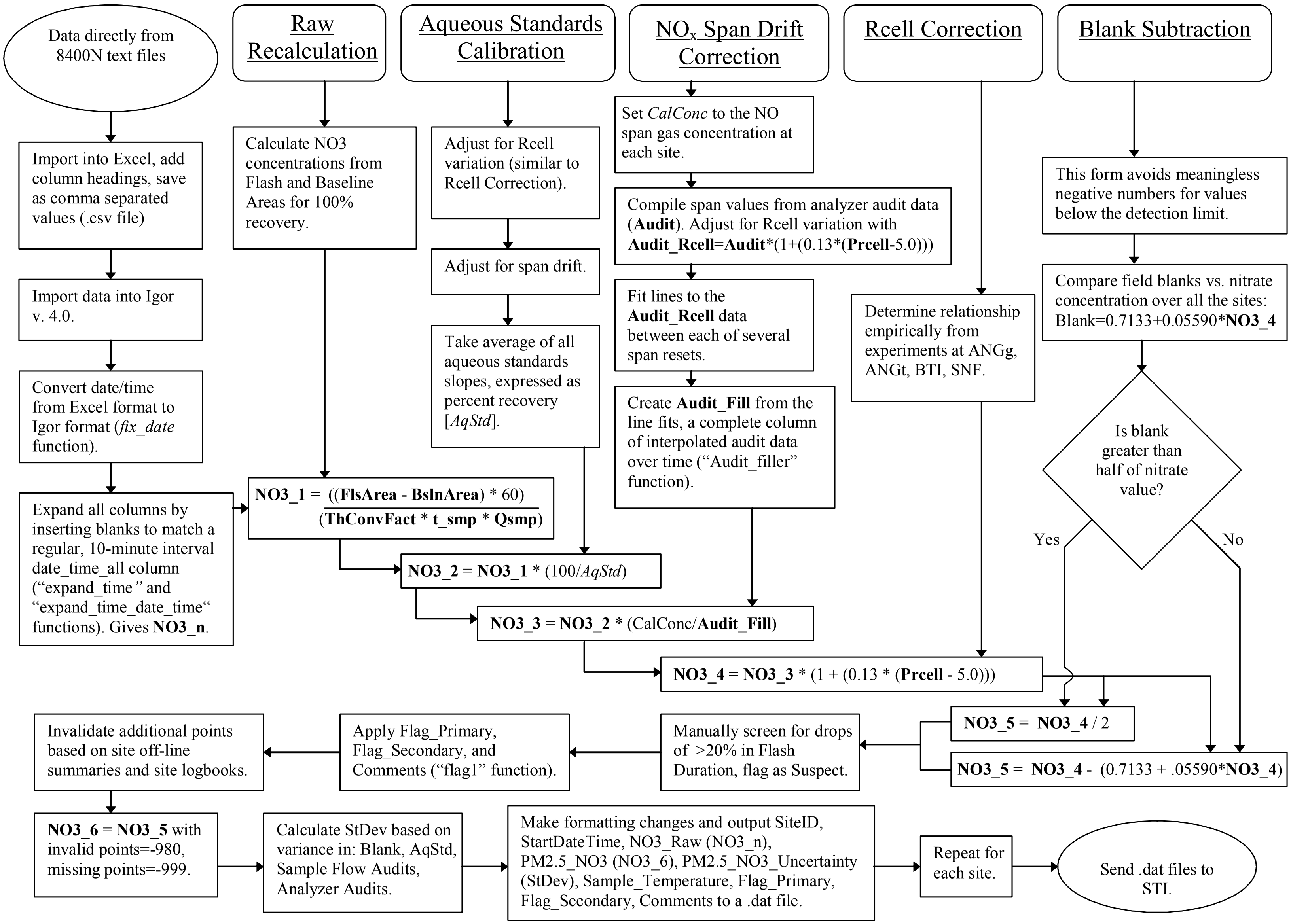


Figure 3-5. Data validation flow chart for 10-minute nitrate measurements (prepared by B. Kirby, ADI).

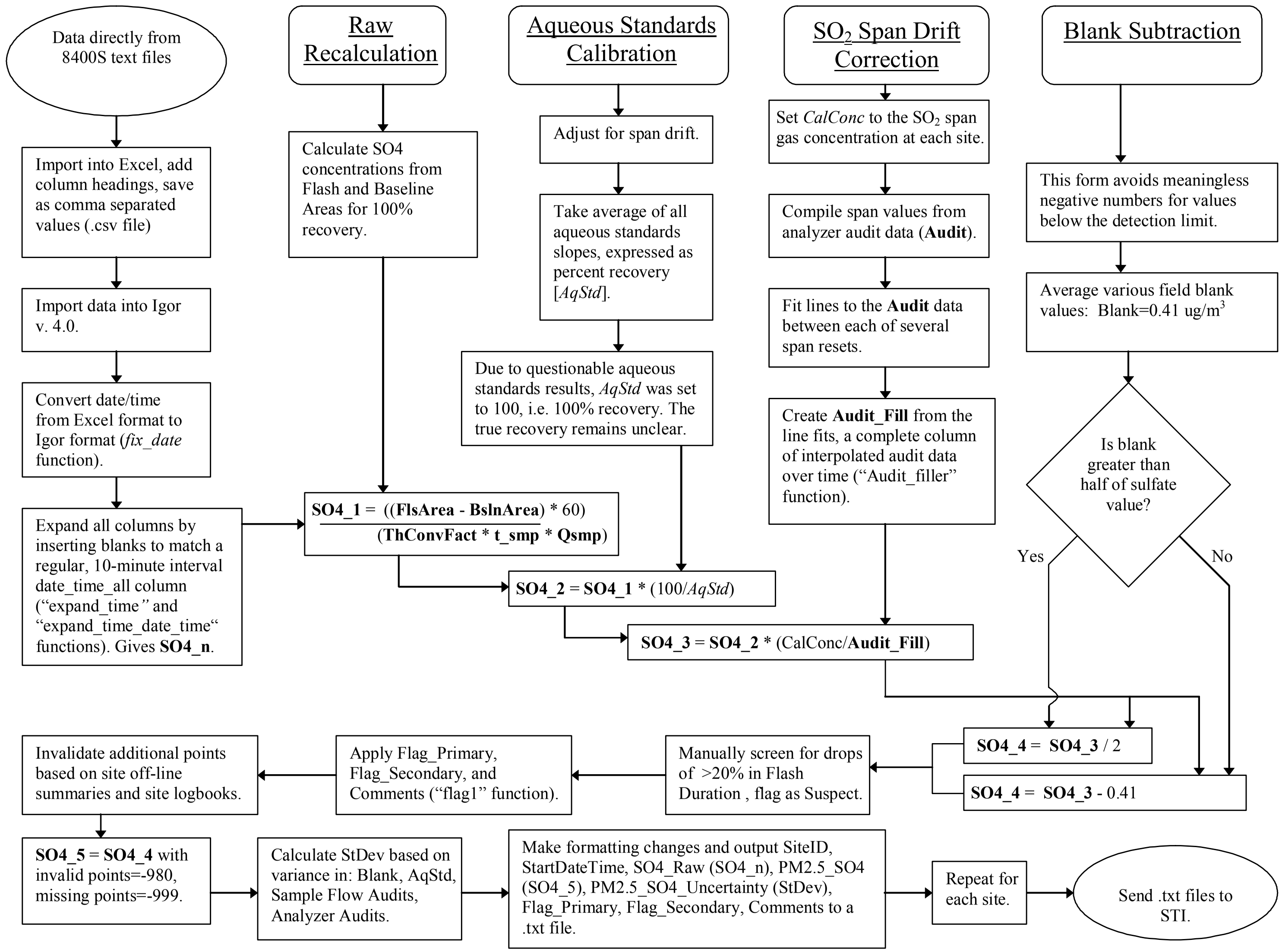


Figure 3-6. Data validation flow chart for 10-minute sulfate measurements (prepared by B. Kirby, ADI).

- Flagging obvious outlier and invalid data points (provided documentation is available to verify instrument problem).
- Extracting and reporting calibration, zero, and span data.
- Screening calibration, zero, and span data for invalid calibration points.
- Calculating calibration slopes, intercepts, and baseline zeros. The measurement experts then review results and determine time periods over which the calibration adjustments apply.
- Applying zero corrections and calibrations.
- Graphically reviewing corrected time series plots (with a focus on any transitions between calibration adjustments).
- Running range checks on corrected data to screen for remaining outliers.
- Annotating all modifications to the data set in an electronic log.

Data that were flagged for reasons such as power failure or a channel down due to calibration, audit, or other reasons, were either not loaded into the final database or were invalidated later. All data loaded into the database were screened for outliers and maximum rates of change between averages. The data management staff verified all the questionable data using daily summaries and site and instrument logbook documentation.

Instrument specific data validation issues are discussed in Section 3.6.

3.3.2 OP Codes

The DAS was used to apply OP codes to the data. The default (valid data) code was 0, an OP code of 7 was instrument error, and an OP code of 9 indicated missing data. All other codes were reserved to relay instrument-specific operation information. OP codes were assigned to gaseous instruments to indicate automated zeroes and spans. For example, when a gas-phase titration (GPT) span calibration on the NO/NO_y instrument was initiated by the DAS, an OP code of 2 was applied to all NO/NO_y data records during the calibration process. Descriptions of the various OP codes used in this study are found in Table 2-2. These OP codes were converted to QC codes using a cross-reference table during the automated import process.

3.3.3 Screening Criteria

With the help of the measurement experts, screening criteria were prepared to assist in data validation. These criteria are provided in **Table 3-1**. Although automatic screening was not implemented during the import process, the data technician followed these guidelines in the daily QC checks of the data. These criteria also provided guidance to data technicians who reviewed the data in later validation steps.

Table 3-1. Example screening criteria.

Instrument	Parameter	Screening Criteria
Aethalometer	BC	Should not drop below -1 nor exceed 20 μg/m^3
		Point-to-point variation should not exceed 1 μg/m^3
		Six consecutive values should not be equal
Aethalometer	Flow	Point-to-point variations should not exceed 0.5 LPM
Nephelometer	b_{sp}	Should not drop below -1 or exceed 2000 Mm-^{1}
		Six consecutive values should not be equal
Nephelometer	RH	Should not drop below 7% or exceed 80%
		Point-to-point variations should not exceed 10%
Nephelometer	Temp	Should not drop below 260 K or exceed 325 K
Ozone	Ozone	Should not drop below -2 or exceed 200 ppb
		Point-to-point variation should not exceed 30 ppb
		Six consecutive values should not be equal. *This test was not applied between 0000 and 0600 hours.*
NO/NO_y	NO/NO_y urban	Should not drop below -1 nor exceed 700 ppb
		Point-to-point variation should not exceed 30 ppb
		NO should not exceed NO_y
		30 consecutive values should not be equal
NO/NO_y	NO/NO_y rural	Should not drop below -1 or exceed 300 ppb
		Point-to-point variation should not exceed 30 ppb
		NO should not exceed NO_y
		30 consecutive values should not be equal
OCEC	OCEC	Should not drop below 0 or exceed 20 μg/m^3
		OC concentrations should not exceed OCEC concentrations
		Three consecutive values should not be equal
BAM	Concentration	PM_{25} should not exceed PM_{10}
BAM	10 μm	Should not drop below -5 or exceed 500 μg/m^3
		Three consecutive values should not be equal
BAM	2.5 μm	Should not drop below -5 or exceed 200 μg/m^3
		Three consecutive values should not be equal

3.4 ASSESSING CALIBRATION DATA

3.4.1 Overview

The approach to data validation, including review and application of calibration data, was similar among all instruments. In reviewing the calibration data, the following steps were generally taken:

1. Assemble and review instrument logbooks. We have found in our fieldwork that each instrument has its own peculiarities. Similarly, each operator approaches his/her job differently. By reviewing the logbooks, the QC technician obtains an overall view of how frequently the instrument was attended to, how much detail the operator put into comments, etc. Site logs were also inspected. For example, did a particular spike in NO concentration correspond to the arrival of the technician at the site?
2. Summarize instrument configuration. The QC technician needed to have an overall view of the instrument, including photographs of components as they were set up in the field.
3. Prepare a list of possible instrument problems. The QC technician used the logbooks and instrument configuration (and sometimes the instrument manual) to anticipate potential problems with the system. For example, if filter changes were frequently made on an instrument, the potential for a leak might be higher than if filter changes were rare.
4. Assess data completeness. A table containing stop and start times for the instrument as well as periods of missing data helped to set the windows of operation.
5. Inspect nightly calibrations and matrix zeroes. The QC technician needed to know the schedule, concentrations used, and other details about the calibrations. In this step, plots of the data were prepared and slopes and intercepts for adjustments were computed.
6. Recommend adjustments to the data. After review of the calibration information, the QC technician worked with the measurement expert to prepare a list of calibration adjustments and the time period over which the adjustments should be applied.

The following sections contain information specific to each instrument with regard to assessing calibration data.

3.4.2 Ozone

Scheduled nightly zero-span, manual zero-span, and multipoint calibration of ozone concentration data were pulled from the 1-minute database and compiled into an MS Excel spreadsheet for visualization using time series plots. An example calibration schedule that provides information about OP code assignments, gases sent to the analyzer, and expected responses is shown in **Table 3-2**. Zero calibration data records, those records with an assigned OP code of 5, were queried in the database and exported to MS Excel for review. Similarly, span calibration information having an OP code of 4 was queried and exported to the MS Excel file. At the beginning of the study, and when new instruments were added, the data collected during the calibrations may not have had assigned OP codes or may have been incorrectly coded. In these cases, calibration data were manually pulled from the database, based on the schedule information; site logs, instrument logs, and other material that detailed expected calibration times; and visual examination of the data.

Table 3-2. Example ozone zero-span calibration check with zero air and 80 ppb ozone.

Clock Time PST	Delta Time (min.)	Comments	Analyzer Affected	Gas Sent to Analyzer	OP Code Assigned	QC Code Assigned	Desired Response (ppb)
0244	0	Start purging the calibration lines	NONE	NONE	NONE	NONE	NONE
0245	1	Send zero air to ozone analyzer	O_3	zero air	5	3	0
0246	2	Check the flow rates	O_3	zero air	5	3	0
0258	14	Check the flow rates	O_3	zero air	5	3	0
0300	16	Start sending 80 ppb ozone	O_3	80 ppb O_3	4	4	80
0301	17	Check the flow rates	O_3	80 ppb O_3	4	4	80
0310	26	Check the flow rates	O_3	80 ppb O_3	4	4	80
0312	28	End calibration and purge calibration lines	O_3	NONE	NONE	4	NONE
0313	29	Ambient recovery	O_3	NONE	NONE	4	NONE
0315	31	Data valid again	NONE	NONE	NONE	NONE	NONE

Once the calibration information was exported to MS Excel, the data technician organized the data and calculated average zero and span readings for each calibration that occurred. Standard deviations and ratios of observed response to expected response were also calculated. The number of data records comprising an average value varied and was dependent on the stability of the readings for each individual instrument. For example, although the ozone calibration gas sent to the Sierra Nevada Foothills ozone instrument during the nightly zero-spans may have occurred over a period of 12 minutes (i.e., 12 1-minute records), only the last 4 or 5 data records were used in calculating the average, reflecting data when gas flow to the instrument and concentrations had reached a more steady state. Time series plots of average zeroes, average spans, zero-adjusted spans (subtracting the most recent average zero from the average span), and efficiencies (unadjusted and zero-adjusted ratios of observed averaged span concentrations to expected span concentrations) were created and designated as originating from automatic zero-spans, manual zero-spans, or multipoint calibrations.

3.4.3 NO/NO_y and Nitric Acid

NO and NO_y concentration data during nightly zero-spans, manual zero-spans, biweekly calibrations, and multipoint calibrations were queried from the 1-minute database according to OP codes assigned during specific instrument checks. Data collected during the zero check were assigned an OP code of 5, the NPN span check an OP code of 3, the NO span check an OP code of 1, the GPT converter check an OP code of 2, the ammonia check an OP code of 4, and the nitric acid check an OP code of 8. A description and schedule for these checks are provided in **Table 3-3**. For calibrations that occurred when OP code information was unavailable, data were queried according to expected times found by consulting various sources or upon visual examination. Data were averaged for each type of zero or span check using the same procedures

Table 3-3. Example NO_y zero-span calibration check with zero air, NPN, NO, and NO_2, followed by zero air.

Clock Time PST	Delta Time (min.)	Comments	Analyzer Affected	Gas Sent to Analyzer	OP Code Assigned	QC Code Assigned	Desired Response (ppb)	
							NO	NO_y
0044	0	Start purging the calibration lines	NONE	NONE	NONE	NONE	NONE	NONE
0045	1	Start sending calibration zero gas to the analyzer	NO_y	zero air	5	3	0	0
0046	2	Check the flow rates	NO_y	zero air	5	3	0	0
0053	9	Check the flow rates	NO_y	zero air	5	3	0	0
0055	11	Start sending NPN at 90 ppb to the analyzer	NO_y	90 ppb NPN	3	4	0	>70
0056	12	Check the flow rates	NO_y	90 ppb NPN	3	4	0	>70
0113	29	Check the flow rates	NO_y	90 ppb NPN	3	4	0	>70
0115	31	Start sending NO at 90 ppb to the analyzer	NO_y	90 ppb NO	1	4	90	90
0116	32	Check the flow rates	NO_y	90 ppb NO	1	4	90	90
0126	42	Check the flow rates	NO_y	90 ppb NO	1	4	90	90
0128	44	Start sending NO at 90 ppb, combined with 60 ppb to create 60 ppb of NO_2 and 30 ppb of NO	NO_y	60 ppb NO_2/ 30 ppb NO	2	4	30	90
0129	45	Check the flow rates	NO_y	60 ppb NO_2/ 30 ppb NO	2	4		
0139	55	Check the flow rates	NO_y	60 ppb NO_2/ 30 ppb NO	2	4	30	90
0141	57	Start sending zero air to the analyzer	NO_y	zero air	5	3	0	0
0142	58	Check the flow rates	NO_y	zero air	5	3	0	0
0149	65	Check the flow rates	NO_y	zero air	5	3	0	0
0151	67	End calibration	NO_y	NONE	NONE	4	NONE	NONE
0152	68	Ambient recovery	NO_y	NONE	NONE	4	NONE	NONE
0200	76	Data valid again	NO_y	NONE	NONE	NONE	NONE	NONE

that were used in preparation of the ozone calibration data described above. Time series plots of average zeroes, average spans, zero-adjusted spans (subtracting the most recent average zero from the average span), and efficiencies (unadjusted and zero-adjusted ratios of observed averaged spans to expected spans) were created for the zero, NO span, NPN span, and GPT span checks.

The nitric acid instrument, almost identical to the NO_y analyzer, collects NO_y and NO_y-HNO_3 data. This instrument was plumbed to, and therefore received, the same calibration information as the NO_y analyzer. NO_y and NO_y-HNO_3 data were pulled from the 1-minute database table according to the same OP code assignments for the NO/NO_y data. The same protocol was followed in averaging and plotting nitric acid calibration data as was used for the NO_y instrument.

3.4.4 Nephelometer

The Level 0 nephelometer data did not contain an OP code to indicate the times during which the instruments were being calibrated. Therefore, these times were determined from the task sheets and instrument logs, and by visually inspecting the data for the characteristic pattern in the b_{sp} and RH data produced by a calibration. The data recorded during calibrations were copied from the 1-minute (analog data) and 5-minute (digital data) data files into an MS Excel spreadsheet for each monitoring site. The 1-minute data were averaged during the times of stable zero and span readings; and when the duration of the calibrations was long enough, the 5-minute data were also averaged. When, as was often the case, a stable zero or span was not maintained for a full 5-minute averaging time, no calibration value was calculated from the digital data. In addition, zero and span values read by the field technician from the nephelometer display were copied from task sheets and entered into the MS Excel spreadsheets.

Either CO_2 or Freon 134a (SUVA) were used for the span calibrations. The nephelometer temperature and ambient pressure (manually set in the nephelometer according to the site elevation) were used in the MS Excel spreadsheet to calculate the light-scattering coefficient of the span gas. Because the nephelometer is calibrated to read zero when filled with particle free air, the correct b_{sp} reading during the span calibration is the scattering coefficient of the span gas minus the scattering coefficient of air. The spreadsheet calculated the ratio of the b_{sp} measured by the nephelometer during the span calibration to the correct value.

For each site, six time series plots were prepared showing the zero and the span ratio for each of the three sources of data (digital, analog, operator). Different symbols were used for CO_2 and SUVA calibrations. Times when the nephelometer calibrations would be expected to change, e.g., because of a recalibration or instrument replacement, were tabulated. An analyst used these time series plots and the underlying data tabulations to estimate the calibrations to be applied to the b_{sp} readings.

Analysis of the calibration and audit data from the nephelometers at the CRPAQS satellite sites by (Richards et al., 2001a; 2001b) showed that after a few outliers were omitted, the mean zero from all calibrations was 0.4 ± 1.4 Mm^{-1}, where the uncertainty is the standard deviation about the mean. The audit data were essentially identical. These same calibration data gave a span ratio of 0.99 ± 0.04 and, again, the audit data were essentially identical. Therefore, when permitted by the calibration data, an effort was made to estimate the zero calibration to within 1 Mm^{-1} and the span ratio to within a few percent. When the calibration data were stable enough to justify the calculation, the MS Excel spreadsheet was used to calculate an average zero and span to be applied.

Calibrations were applied to the data only when the zero was off by more than 2 Mm^{-1} and the span ratio was off by more than 5%, or 0.05. It is expected that dust accumulation on the walls of the scattering chamber and dark trap will cause the nephelometer zero reading to increase with time. Therefore, for some time intervals, a zero calibration that changed linearly with time was applied to the data. **Table 3-4** lists the calibration factors that were applied to the data.

Table 3-4. Calibration factors applied to the nephelometer data by site and instrument serial number (S/N).

Site	Neph S/N	Start Date	End Date	Zero From Mm^{-1}	Zero To Mm^{-1}	Span
Angiola	262	5/9/2000	10/17/2000	-1	3.2	0.96
Angiola	262	10/17/2000	1/3/2001	3.2	5.5	1.02
Angiola	262	1/3/2001	2/7/2001	5.5		1.08
Angiola 1 meter	194	12/14/2000	2/7/2001	4.64		1.03
Angiola 50 meter	192	8/18/2000	1/11/2001	1.5		1.00
Angiola 50 meter	192	1/19/2001	2/8/2001	2.6		1.00
Angiola 100 meter	193	8/18/2000	1/11/2001	1.8		0.95
Angiola 100 meter	192	1/11/2001	1/19/2001	2		1.00
Angiola 100 meter	193	1/19/2001	2/8/2001	-0.51		0.95
Bakersfield	213	3/20/2000	6/20/2000	0.5	3.6	0.95
Bakersfield	213	6/20/2000	8/16/2000	1.8	3.7	0.98
Bakersfield	194	10/18/2000	11/15/2000	1.20	2.80	0.99
San Jose	228	1/22/2000	9/8/2000	-0.89	0.84	0.88
San Jose	228	9/8/2000	2/8/2001	0.84	2.02	0.94
Walnut Grove Tower	278	11/26/2000	12/14/2000	2.7		1.13

3.4.5 BAM

Only flow rate calibrations are required for the BAM instruments. The BAM-indicated flows were compared to actual flows as measured by a transfer standard. Repeated discrepancies of >10% between the indicated and actual flow rates warranted adjustments to the flow and concentration data. Flow data at any of the sites with significant flow differences were adjusted by the ratio of the actual to the indicated flow rates. Concentration adjustments were subsequently applied to the same dataset using the inverse of the applied flow correction. These changes were noted in the database delivered to ARB.

3.4.6 Aethalometers

Flow audits were performed on the Aethalometer using a BIOS flow meter. If the volumetric flow differed from the set point flow rate (6.9 LPM) by more than 10%, corrective action was attempted. If corrective action was not completed successfully, data through the prior successful flow audit were invalidated. If the instrument-indicated flow rate differed significantly from the volumetric flow, the flow standard was recalibrated and the audit repeated.

Optical test strips were also used to verify that the Aethalometer's lamp performance remained consistent over its operating period. These strips performed well on all CRPAQS Aethalometers.

In addition, dynamic zero checks were performed on the Aethalometer to confirm that its baseline remained at zero; these zeroes routinely passed.

3.4.7 OCEC

Carbon dioxide calibrations were performed monthly on the OCEC instrument to determine the accuracy of its LiCor sensor. Three different concentrations of CO_2 gas were used to perform these audits, and the following criteria were applied:

- Zero air (0 ppm) +/- 25 ppm
- CO2-Low (%) +/- 5% of actual tank concentration (~405 ppm)
- CO2-High (%) +/- 5% of actual tank concentration (~2401 ppm)

The concentrations did not always fall within the acceptable ranges but were typically within 10 to 15%. There is no indication in the documentation that a complete LiCor sensor calibration (or adjustment) was ever performed. Neither flow nor concentration factors were applied to the data. If calibration results did not fall within the acceptable range, affected data and potentially affected data were invalidated after discussion with the measurement expert.

Flow audits were performed using a BIOS flow meter. If the volumetric flow differed from the set point flow rate (16.67 LPM) by more than ±0.5 LPM, corrective action was attempted. If corrective action was not completed successfully, data since the prior successful flow audit were invalidated. If the instrument-indicated flow rate differed significantly from the volumetric flow, the flow standard was recalibrated.

3.4.8 SO_2

The nightly zero-span checks that were performed were useful in identifying operating tendencies of the SO_2 monitor. These checks showed the monitor to be stable throughout the monitoring period with a slope of 1.02 to 1.04. These slopes would slightly overcorrect the data since adequate time to reach a stable span value was not achieved.

The results of multipoint calibrations or stable value manual zero-spans were reviewed. These results were used to determine correction factors that were applied to the data (see **Table 3-5**).

Table 3-5. Correction factors for SO_2.

Site	Date/Time Range	Slope
Bakersfield	10/20/00 – 1/5/01 0259 PST	1.01
Bakersfield	1/5/01 0300 PST to end	1.03

3.4.9 PAN/NO_2

Calibration of the NO_2 GC was performed using the corrected NO values from the daily calibration of the NO_y analyzer. Daily calibrations of the NO_y analyzer included two modes useful for the NO_2-GC calibration: the NO span check and the GPT converter check. The concentration of NO_2 during the GPT span check was determined by taking the difference in the corrected analyzer NO response between the NO span check and the GPT span check, assuming the NO had been converted to NO_2. A calibration factor for the NO_2 GC was then calculated as the NO_2 GC's response during the GPT check divided by the calculated NO_2 concentration.

These daily calibration factors were linearly interpolated for each minute of the monitoring period and applied to the data in a manner similar to that described for ozone data.

PAN calibration was performed using the PAN/NO_2 GC's relative response ratio calculated from the PAN and NO_2 response as (counts per peak height/ppb PAN) divided by (counts per peak height/ppb NO_2). This response ratio was applied to the NO_2 calibration factor to determine the PAN calibration factor. This PAN calibration factor was then multiplied by the peak magnitude to determine PAN concentration.

3.4.10 Nitrate and Sulfate

Span checks were performed every other day for both the sulfate (using 1 ppm SO_2 gas) and nitrate (using 5 ppm NO gas) instruments. Semi-monthly calibrations were performed using aqueous standards. Nitrate and sulfate concentrations were corrected for span drift and percent recovery (based on aqueous standards) as discussed in Section 3.6.

3.4.11 Climet OPC and PMS Lasair OPC

The calibration of these instruments consisted of a flow check performed at the inlet located on the sampling trailer's roof, flow checks performed on individual instruments, dynamic zeroes, and polystyrene latex (PSL) checks. These tasks were scheduled to be performed monthly but were supposed to be performed more frequently during the winter intensive periods. Although these tasks were sometimes not completed for months at a time, data were considered valid as long as the checks passed and no other problems were documented.

Roof inlet flow checks were performed using a calibrated flow standard, such as the BIOS or Gilibrator. Flows within 10% of the set point (which varied by inlet) were considered valid. Flows outside this range were grounds for suspecting the two largest Climet OPC bins. All data from the two largest Climet OPC bins were flagged with a QC code of 1 ("data wholly or partially compromised") because, even under normal conditions, some particles in that size range are removed by the inlet (Kreisberg, 2002).

Individual instrument flows were also checked with the BIOS or Gilibrator; different models accommodate different flow ranges. Instrument flow checks were considered valid if the flow standard measured within 10% of the flow rate set point *and* if the instrument-indicated flow rate agreed within approximately 10% of the flow standard. The Climet OPC instruments passed all flow checks. The PMS Lasair OPC, however, suffered many flow problems.

- In July 2000, its indicated sample volume dropped steadily from approximately 0.029 L to 0.020 L. Flow checks in late July indicated that the flow was well below the set point but that it was not quite as low as the Lasair OPC indicated. Unfortunately, no other flow checks were performed until January 9, 2001, with the exception of an independent flow audit on October 23, 2000.
- The October 23, 2000, audit showed agreement between the indicated flow and the flow standard, but the indicated flow was approximately 50% higher than the set-point flow. The indicated flow rate had been very high since late October when the Lasair OPC was returned to Angiola for the second time.
- The incorrect flow may be partially explained by the fact that, between the end of July and the end of October, the instrument was twice shipped back to PMS, where its altitude setting was significantly adjusted (from a correct value of 443 feet to 5850 feet). Because the Lasair OPC's flow is controlled by a differential pressure sensor (across a critical orifice) and not by a mass flow meter, this adjustment affects the flow rate. However, the effect of adjusting the altitude setting to a higher altitude should have had the opposite effect; flow should have been reduced because the sensor would have detected more-than-adequate pressure. Moreover, the flow discrepancy was not proportional to the change in altitude. It is possible that the sensitive altitude/flow adjustment knob may have been jostled in transport. Until the altitude setting and flow rate problems were corrected, all Lasair OPC data were flagged as suspect or invalid.

Dynamic zeroes were performed by attaching a high efficiency particulate air (HEPA) filter to each instrument's inlet. The zero was valid if, after 5 minutes, the instrument read a total count of 10 or fewer particles over a minute or so. The PMS Lasair OPC consistently passed the dynamic zeroes. The Climet OPCs, however, did not; all three Climet OPC instruments failed the zeroes at various points. Because the high zero counts were typically in the smallest bin, this bin was suspect whenever a dynamic zero check failed. The failed zeroes of at least one Climet OPC were due in part to a misaligned laser diode.

PSL checks were performed on the Climet OPCs, the PMS Lasair OPC, and the SMPS. For the PSL checks, moderate concentrations of PSL spheres of known diameter were nebulized and injected in a diluted sample stream. Five different size spheres were used: in diameter, 4.6 µm, 1.4 µm, 0.89 µm, 0.58 µm, and 0.23 µm. The largest four spheres were measured by the Climet OPC; the smaller four spheres were measured by the Lasair OPC, and the smallest spheres were measured also by the SMPS. The operator verified that the majority of the spheres fell into the correct size bin of a given instrument by recording the counts in the relevant bins and/or by compiling the computer-generated printouts of the bin counts. All of the documented

PSL checks on the Lasair OPC, ground Climet OPC, and 100-m Climet OPC passed. The few failed PSL checks on the 50-m Climet OPC were attributed to a misaligned laser diode, as mentioned above in reference to the failed dynamic zeroes.

3.4.12 SMPS

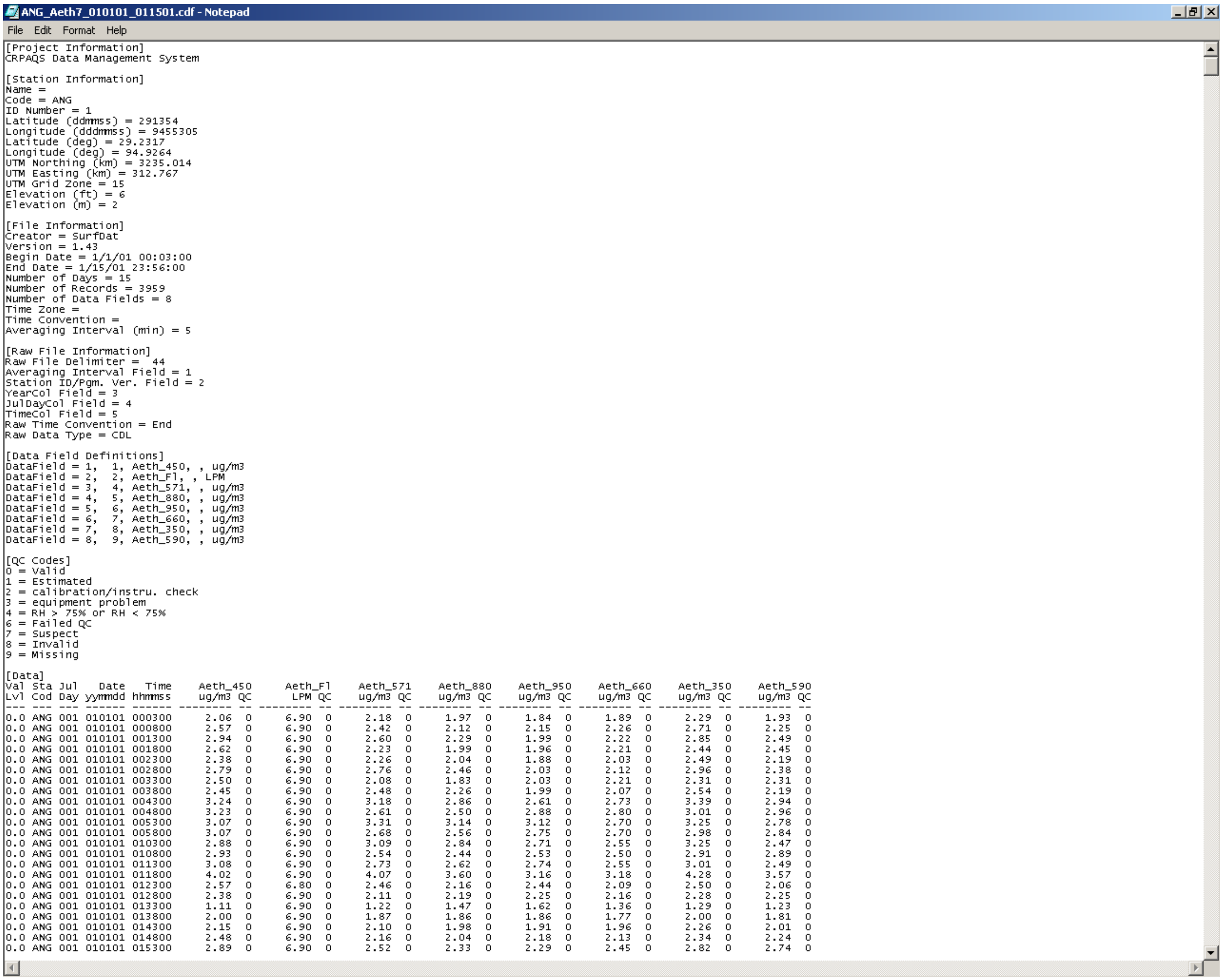
As a check on overall SMPS sizing performance, and specifically of the sheath flow post-process calibration, the PSL data were reprocessed at a finer size resolution of 64 channels per decade. According to the site log and task sheets, thirteen PSL checks were performed using 0.27 μm spheres. Of these, SMPS scans with identifiable peaks could only be found for eleven checks. This left 34 useable samples, the placement of the singly-charged peak of all but one of which matched 0.27 μm to within approximately 5%. A second scan corresponding to the same check as the one outlier was within the 5% range. Three PSL checks were performed at significantly elevated sheath flow rates - 8.92, 9.43 and 9.61 L/min. For all of these checks, the fit of the measured size to 0.27 μm was significantly better using the corrected sheath flow than using the nominal sheath flow (7 L/min).

3.5 DATA VALIDATION USING SURFDAT

3.5.1 SurfDat and Data Validation Overview

The SurfDat program, created by STI, was used to perform data validation tasks for various parameters. SurfDat allows the user to review data, apply calibrations, invalidate data, and change QC codes. All changes made are automatically logged in a log file for traceability. Several versions of the program were used over the course of the study; however, data validated in earlier software versions maintained the same level of integrity as data validated with later versions. Data were exported as a SurfDat (CDF) file from the CRPAQS data management system, using a program that specifically formatted the selected data into a SurfDat-readable file. The CDF file (e.g., **Figure 3-7**) contained summary information on the site in addition to parameter information and the selected data records with accompanying QC codes. In SurfDat, data could be viewed in tabular (**Figure 3-8**) or graphical (**Figure 3-9**) formats.

Plots were created and viewed for time scales ranging from a minimum of the parameter's sampling frequency to a maximum time spanning the entire data file. In this user-defined way, the QC technician could view the data closely, point to point, or over an expanded period of weeks or months. Data scales were also easily changed to accommodate the desired viewing resolution. A built-in zoom function allowed the technician to quickly highlight areas of interest. Only two parameters at a time could be displayed on a single SurfDat plot, although multiple plots could be created and displayed in a SurfDat window. The limitations made it sometimes necessary to use web site plots and data exported and plotted in MS Excel for comparison among different parameters. SurfDat also provided the QC technician the option to link changes from one parameter to several others (e.g., for the particle-sizing instruments, all size cuts could have QC code changes made even though only two size cuts were plotted on screen).




```
[Project Information]
CRPAQS Data Management System

[Station Information]
Name =
Code = ANG
ID Number = 1
Latitude (ddmmss) = 291354
Longitude (dddmmss) = 9455305
Latitude (deg) = 29.2317
Longitude (deg) = 94.9264
UTM Northing (km) = 3235.014
UTM Easting (km) = 312.767
UTM Grid Zone = 15
Elevation (ft) = 6
Elevation (m) = 2

[File Information]
Creator = SurfDat
Version = 1.43
Begin Date = 1/1/01 00:03:00
End Date = 1/15/01 23:56:00
Number of Days = 15
Number of Records = 3959
Number of Data Fields = 8
Time Zone =
Time Convention =
Averaging Interval (min) = 5

[Raw File Information]
Raw File Delimiter =  44
Averaging Interval Field = 1
Station ID/Pgm. Ver. Field = 2
YearCol Field = 3
JulDayCol Field = 4
TimeCol Field = 5
Raw Time Convention = End
Raw Data Type = CDL

[Data Field Definitions]
DataField = 1,  1, Aeth_450, , ug/m3
DataField = 2,  2, Aeth_Fl, , LPM
DataField = 3,  4, Aeth_571, , ug/m3
DataField = 4,  5, Aeth_880, , ug/m3
DataField = 5,  6, Aeth_950, , ug/m3
DataField = 6,  7, Aeth_660, , ug/m3
DataField = 7,  8, Aeth_350, , ug/m3
DataField = 8,  9, Aeth_590, , ug/m3

[QC Codes]
0 = Valid
1 = Estimated
2 = calibration/instru. check
3 = equipment problem
4 = RH > 75% or RH < 75%
6 = Failed QC
7 = Suspect
8 = Invalid
9 = Missing

[Data]
Val Sta Jul   Date   Time    Aeth_450       Aeth_Fl      Aeth_571      Aeth_880      Aeth_950      Aeth_660      Aeth_350      Aeth_590
Lvl Cod Day yymmdd hhmmss    ug/m3 QC        LPM QC      ug/m3 QC      ug/m3 QC      ug/m3 QC      ug/m3 QC      ug/m3 QC      ug/m3 QC
--- --- --- ------ ------ -------- -- -------- -- -------- -- -------- -- -------- -- -------- -- -------- -- -------- --
0.0 ANG 001 010101 000300     2.06  0     6.90  0     2.18  0     1.97  0     1.84  0     1.89  0     2.29  0     1.93  0
0.0 ANG 001 010101 000800     2.57  0     6.90  0     2.42  0     2.12  0     2.15  0     2.26  0     2.71  0     2.25  0
0.0 ANG 001 010101 001300     2.94  0     6.90  0     2.60  0     2.29  0     1.99  0     2.22  0     2.85  0     2.49  0
0.0 ANG 001 010101 001800     2.62  0     6.90  0     2.23  0     1.99  0     1.96  0     2.21  0     2.44  0     2.45  0
0.0 ANG 001 010101 002300     2.38  0     6.90  0     2.26  0     2.04  0     1.88  0     2.03  0     2.49  0     2.19  0
0.0 ANG 001 010101 002800     2.79  0     6.90  0     2.76  0     2.46  0     2.03  0     2.12  0     2.96  0     2.38  0
0.0 ANG 001 010101 003300     2.50  0     6.90  0     2.08  0     1.83  0     2.03  0     2.21  0     2.31  0     2.31  0
0.0 ANG 001 010101 003800     2.45  0     6.90  0     2.48  0     2.26  0     1.99  0     2.07  0     2.54  0     2.19  0
0.0 ANG 001 010101 004300     3.24  0     6.90  0     3.18  0     2.86  0     2.61  0     2.73  0     3.39  0     2.94  0
0.0 ANG 001 010101 004800     3.23  0     6.90  0     2.61  0     2.50  0     2.88  0     2.80  0     3.01  0     2.96  0
0.0 ANG 001 010101 005300     3.07  0     6.90  0     3.31  0     3.14  0     3.12  0     2.70  0     3.25  0     2.78  0
0.0 ANG 001 010101 005800     3.07  0     6.90  0     2.68  0     2.56  0     2.75  0     2.70  0     2.98  0     2.84  0
0.0 ANG 001 010101 010300     2.88  0     6.90  0     3.09  0     2.84  0     2.71  0     2.55  0     3.25  0     2.47  0
0.0 ANG 001 010101 010800     2.93  0     6.90  0     2.54  0     2.44  0     2.53  0     2.50  0     2.91  0     2.89  0
0.0 ANG 001 010101 011300     3.08  0     6.90  0     2.73  0     2.62  0     2.74  0     2.55  0     3.01  0     2.49  0
0.0 ANG 001 010101 011800     4.02  0     6.90  0     4.07  0     3.60  0     3.16  0     3.18  0     4.28  0     3.57  0
0.0 ANG 001 010101 012300     2.57  0     6.80  0     2.46  0     2.16  0     2.44  0     2.09  0     2.50  0     2.06  0
0.0 ANG 001 010101 012800     2.38  0     6.90  0     2.11  0     2.19  0     2.25  0     2.16  0     2.28  0     2.25  0
0.0 ANG 001 010101 013300     1.11  0     6.90  0     1.22  0     1.47  0     1.62  0     1.36  0     1.29  0     1.23  0
0.0 ANG 001 010101 013800     2.00  0     6.90  0     1.87  0     1.86  0     1.86  0     1.77  0     2.00  0     1.81  0
0.0 ANG 001 010101 014300     2.15  0     6.90  0     2.10  0     1.98  0     1.91  0     1.96  0     2.26  0     2.01  0
0.0 ANG 001 010101 014800     2.48  0     6.90  0     2.16  0     2.04  0     2.18  0     2.13  0     2.34  0     2.24  0
0.0 ANG 001 010101 015300     2.89  0     6.90  0     2.52  0     2.33  0     2.29  0     2.45  0     2.82  0     2.74  0
```

Figure 3-7. Screenshot of CDF file format.

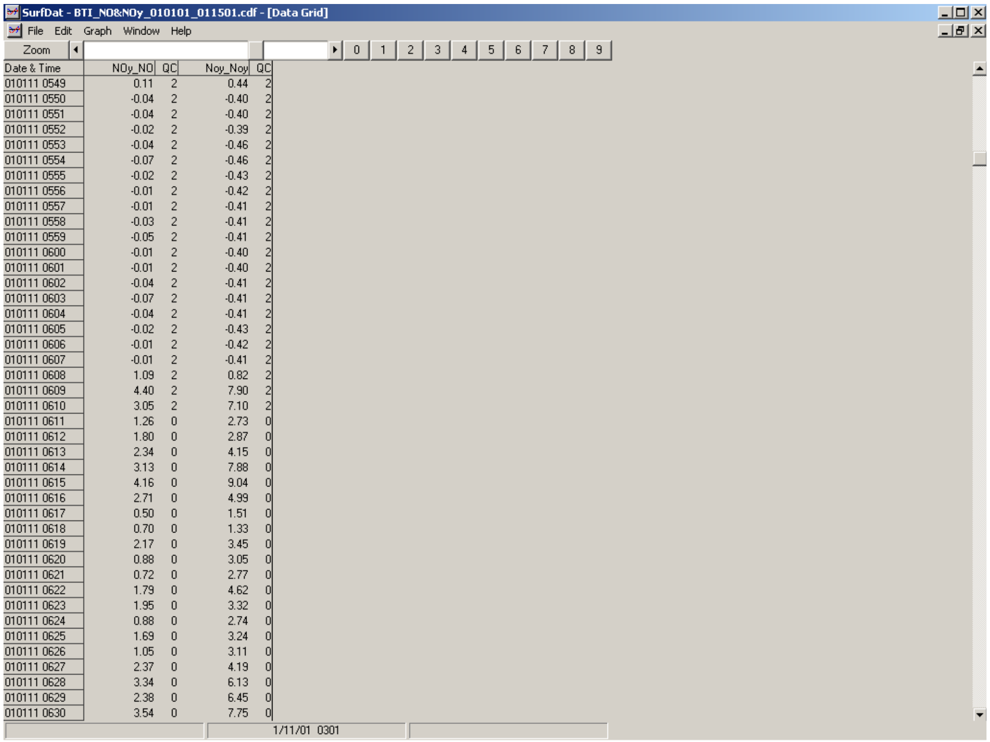


Figure 3-8. Screenshot of SurfDat data grid.

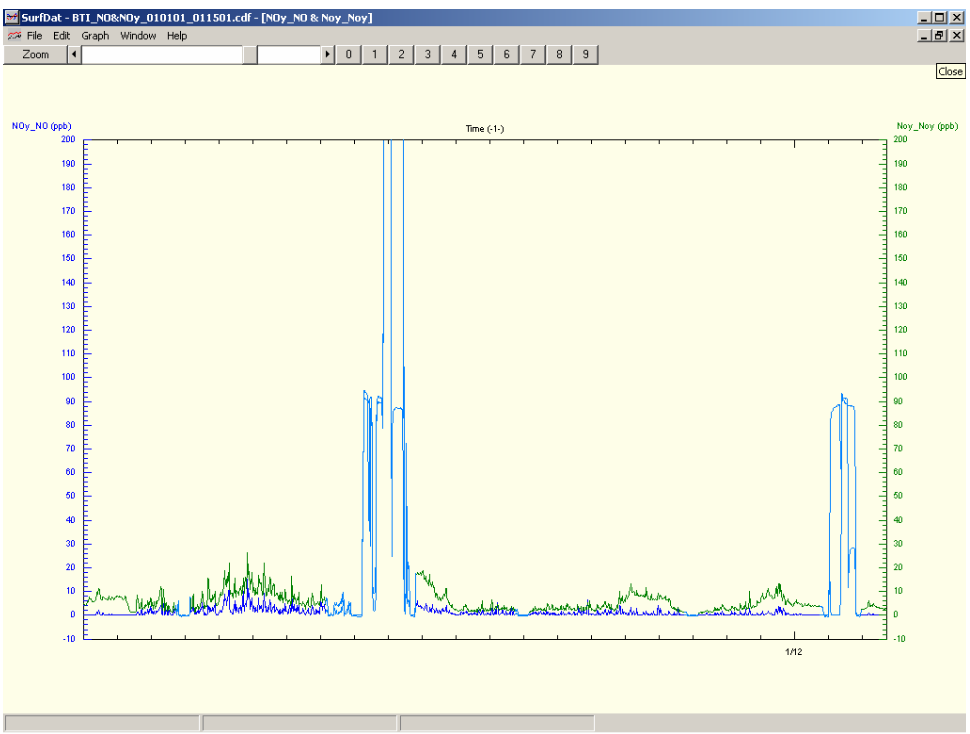


Figure 3-9. Screenshot of SurfDat time series plot.

The data were validated using two different methods within the SurfDat program. Applying offsets and slopes for large sections of data or for identifying off-line periods with sharply defined start/end times was performed with a batch-editing tool (**Figure 3-10**). Otherwise, the QC technician viewed SurfDat data plots over select time periods and manually edited the data by directly selecting the affected areas on the graph. Data QC with the batch-edit method was useful in flagging periods with definitive start and end times. For example, invalid periods of data with defined start and end times, as were recorded in the instrument off-line summary sheets, could be entered into the batch-edit form quickly and effectively. Editing the data in batches was also useful for flagging long time periods with a single flag or during very short time periods. Manually flagging data directly on the plots, although more convenient, was a much less accurate way of flagging data, requiring a steady and patient hand in selecting desired data points. However, the visual advantage of flagging the data directly on the plots accounted for fewer validation errors in the long run, as the outcome of flagging with the batch-edit method could not be seen immediately. Flagging for outliers, suspect data, and other problems not specifically identified in off-line summaries and associated instrument or site logs, was performed by application directly on the time series plots. Both methods of validation prompted the user to enter a comment (to provide the rationale) for every change made. The action taken and associated comments were recorded in log files; these logs were subsequently compiled and exported as notes accompanying the final deliverable data records.

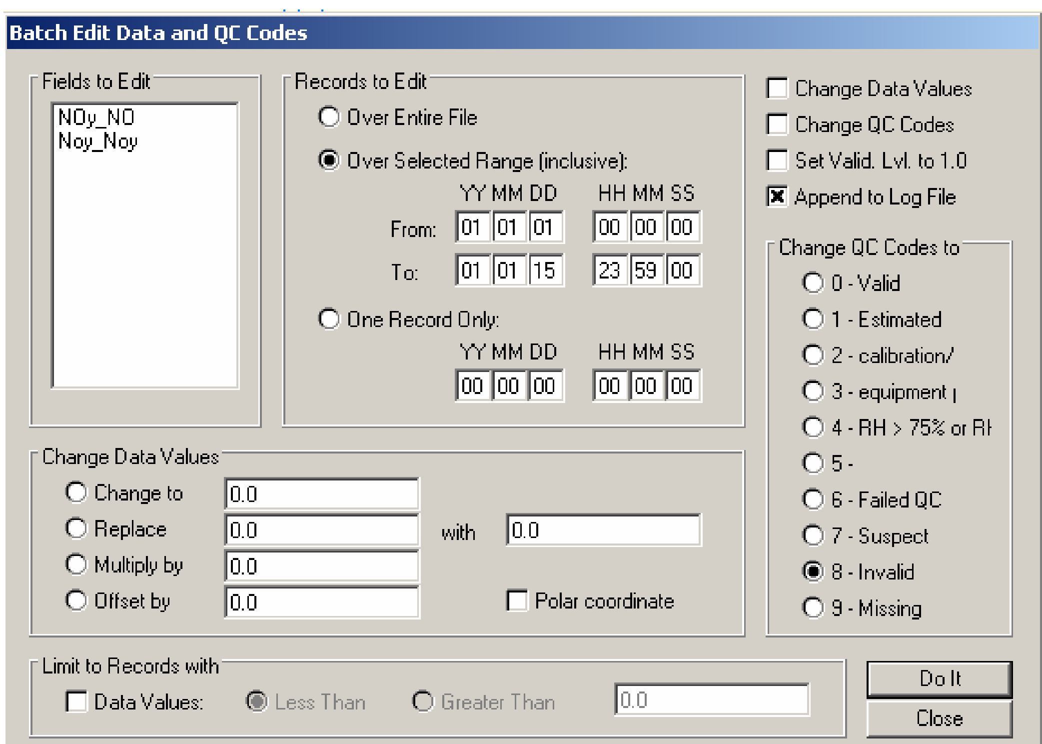


Figure 3-10. Screen shot showing batch-editing of SurfDat file.

Following data validation and application of calibrations, the CDF file and accompanying log file were transferred from SurfDat to the SQL database. Modified data values and QC codes were updated in the SQL databases from which they originated; however, the RawValue associated with each record was unchanged. The accompanying log files were parsed and stored in the LogData database (**Figure 3-11**).

The data validation flags used in this project are shown in **Table 3-6**.

Table 3-6. QC flags and definitions used by STI in this project.

Flag	Definition
0	valid
1	estimated
2	calibration
3	equipment problem/instrument failure
4	RH>75% or RH <7% (nephelometer)
5	tape transfer problem (Aethalometer)
6	failed QC
7	suspect
8	invalid
9	Missing
10	averaged 1-minute data (nephelometer)[a]
11	slightly suspect, heater on continuously (nephelometer)

[a] Analog (1-minute data) were used when the digital 5-minute data were unavailable. Data deemed valid.

3.5.2 Exporting Data to SurfDat

Use of SurfDat requires SurfDat-formatted files that can be output from the main CRPAQS database at STI. SurfDat files were created in the CRPAQS data management system using a program created especially for this task. **Figure 3-12** shows the form for creating these files. A SurfDat file must contain information from only one site but can contain any number of parameters, as long as they are of the same sampling interval. Therefore a 1-minute resolution file, containing any 1-minute data parameters (such as NO_y and ozone), could be created, but a file containing a 5-minute parameter (e.g., Aethalometer BC) and a 60-minute parameter (e.g., BAM $PM_{2.5}$ mass) could not. Data typically combined for viewing in SurfDat included nephelometer b_{sp}, ozone, and NO_y; ozone, NO, and NO_y; NO_y from the NO/NO_y instrument and NO_y from the nitric acid instrument; and OPC and b_{sp} data.

Microsoft Access - [DataLog : Table]

File Edit View Insert Format Records Tools Window Help

DataLo	SiteID	ParamID	StartTime	EndTime	Description	Comment	PgmNameAn	Category	Created
44	1	90	3/16/2000 8:30:00 AM	3/16/2000 10:18:00 AM	QC code set to 2	Span from 830-1018 (Dennis) to check O3 output	SurfDat 1.37	change QC cc	1 12:56:47 PM
45	1	90	3/16/2000 10:19:00 AM	3/16/2000 10:23:00 AM	QC code set to 2	ambient recovery following span	SurfDat 1.37	change QC cc	01 1:02:36 PM
46	1	90	3/16/2000 11:00:00 AM	3/16/2000 11:00:00 AM	QC code set to 8 & data value	negative spike; [O3] drops from 34 ppb (10:59) to -8 (1	SurfDat 1.37	change QC &	01 1:11:44 PM
47	1	90	3/21/2000 11:25:00 AM	3/21/2000 11:41:00 AM	QC code set to 2	filter change; valve check; inlet check	SurfDat 1.37	change QC cc	01 3:38:23 PM
48	1	90	3/21/2000 11:42:00 AM	3/21/2000 11:46:00 AM	QC code set to 2	ambient recovery following filter change; inlet and valve	SurfDat 1.37	change QC cc	01 3:45:00 PM
49	1	90	3/28/2000 10:38:00 AM	3/28/2000 10:41:00 AM	QC code set to 2	filter change (1038-1041)	SurfDat 1.37	change QC cc	01 5:30:10 PM
50	1	90	3/28/2000 10:42:00 AM	3/28/2000 10:45:00 AM	QC code set to 2	inlet/fittings checked	SurfDat 1.37	change QC cc	01 5:31:27 PM
51	1	90	3/16/2000 2:45:00 AM	3/16/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:08:50 AM
52	1	90	3/17/2000 2:45:00 AM	3/17/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:09:08 AM
53	1	90	3/18/2000 2:45:00 AM	3/18/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:10:14 AM
54	1	90	3/19/2000 2:45:00 AM	3/19/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:10:22 AM
55	1	90	3/20/2000 2:45:00 AM	3/20/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:10:29 AM
56	1	90	3/21/2000 2:45:00 AM	3/21/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:10:36 AM
57	1	90	3/22/2000 2:45:00 AM	3/22/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:10:43 AM
58	1	90	3/23/2000 2:45:00 AM	3/23/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:10:50 AM
59	1	90	3/24/2000 2:45:00 AM	3/24/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:10:56 AM
60	1	90	3/25/2000 2:45:00 AM	3/25/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:11:02 AM
61	1	90	3/26/2000 2:45:00 AM	3/26/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:11:09 AM
62	1	90	3/27/2000 2:45:00 AM	3/27/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:11:16 AM
63	1	90	3/28/2000 2:45:00 AM	3/28/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:11:24 AM
64	1	90	3/29/2000 2:45:00 AM	3/29/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:11:31 AM
65	1	90	3/30/2000 2:45:00 AM	3/30/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:11:38 AM
66	1	90	3/31/2000 2:45:00 AM	3/31/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:11:44 AM
67	1	90	3/31/2000 3:12:00 AM	3/31/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:13:43 AM
68	1	90	3/30/2000 3:12:00 AM	3/30/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:14:50 AM
69	1	90	3/29/2000 3:12:00 AM	3/29/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:14:59 AM
70	1	90	3/28/2000 3:12:00 AM	3/28/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:15:05 AM
71	1	90	3/27/2000 3:12:00 AM	3/27/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:15:13 AM
72	1	90	3/26/2000 3:12:00 AM	3/26/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:15:21 AM
73	1	90	3/25/2000 3:12:00 AM	3/25/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:15:27 AM
74	1	90	3/24/2000 3:12:00 AM	3/24/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:15:34 AM
75	1	90	3/23/2000 3:12:00 AM	3/23/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:15:40 AM
76	1	90	3/22/2000 3:12:00 AM	3/22/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:15:46 AM
77	1	90	3/21/2000 3:12:00 AM	3/21/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:15:54 AM
78	1	90	3/20/2000 3:12:00 AM	3/20/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:16:00 AM
79	1	90	3/19/2000 3:12:00 AM	3/19/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:16:06 AM
80	1	90	3/18/2000 3:12:00 AM	3/18/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:16:12 AM
81	1	90	3/17/2000 3:12:00 AM	3/17/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:16:18 AM
82	1	90	3/16/2000 3:12:00 AM	3/16/2000 3:18:00 AM	QC code set to 2	ambient recovery following nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:16:23 AM
83	1	90	4/4/2000 10:12:00 AM	4/4/2000 10:17:00 AM	QC code set to 2	filter change	SurfDat 1.38	change QC cc	01 2:45:44 PM
84	1	90	4/4/2000 10:18:00 AM	4/4/2000 10:22:00 AM	QC code set to 2	ambient recovery following filter change	SurfDat 1.38	change QC cc	01 2:46:17 PM
85	1	90	4/11/2000 12:40:00 PM	4/11/2000 12:50:00 PM	QC code set to 2	filter change	SurfDat 1.38	change QC cc	01 3:27:40 PM
86	1	90	4/11/2000 12:51:00 PM	4/11/2000 12:55:00 PM	QC code set to 2	ambient recovery following filter change	SurfDat 1.38	change QC cc	01 3:29:01 PM
87	1	90	4/1/2000 2:45:00 AM	4/1/2000 3:11:00 AM	QC code set to 2	nightly zero-span cal	SurfDat 1.38	change QC cc	1 11:48:50 AM

Figure 3-11. Example of data validation (SurfDat) log comments as stored in the database.

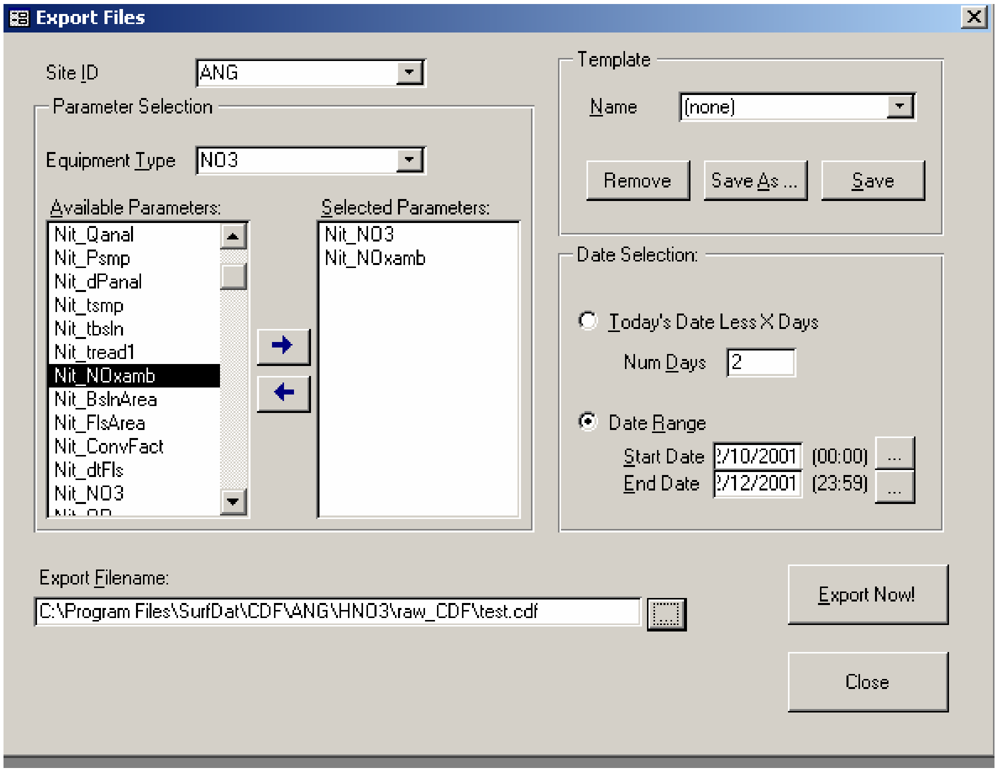


Figure 3-12. Creation of CDF file for export from the SQL database to SurfDat.

One problem with combining data sets occurred when two parameters with the same sampling interval were not synchronized in the actual sampling times. In such a case, SurfDat would create pseudo "zero" value data records. For example, if BAM PM_{10} mass concentration data were sampled on the hour, and BAM $PM_{2.5}$ mass concentration values were recorded 5 minutes past the hour, "zero" data would appear for the BAM $PM_{2.5}$ mass concentration on the hour and for the BAM PM_{10} mass concentration at 5 minutes past the hour. We subsequently rolled back the offset data to the top of the hour to facilitate the comparison.

SurfDat also had file-size limitations. We found that only about 15 days of 1-minute data could be included in a CDF file in order for SurfDat to operate efficiently. For this reason, 1-minute data were routinely divided into two bimonthly files, and 5-minute and 60-minute data were combined by month for each site.

3.5.3 Applying Calibration Data

Once the calibration factors were reviewed by the measurement expert and given final approval, they were applied to the data using the batch-editing tool in SurfDat. Zero offsets were always introduced before slope factors were applied. The data QC technician was prompted to enter a comment for the log file when a change was made; these comments, indicating the applied zero offset or slope value, were tied to every affected data point in the final exported data. During this process, only the data values were changed; QC codes remained the same.

Table 3-7 shows, by site and parameter, the applied calibration factors (zero offsets and slopes) for the gaseous pollutant data.

Table 3-7. Calibrations applied to the gaseous pollutant data.

Site	Instrument	Parameter	Start Date/Time PST	End Date/Time PST	Applied Calibration Factors	
					Zero Offset	Slope
ANGI	HNO_3	NO_y	12/13/00 0000	1/24/01 2359	-0.51	1.02
ANGI	HNO_3	NO_y	1/25/01 0000	2/4/01 2359	-0.51	1.11
ANGI	NO_y	NO	2/1/00 0000	2/24/00 1254	-0.1	1.02
ANGI	NO_y	NO	2/24/00 1255	2/26/00 1227	-0.1	0.79
ANGI	NO_y	NO	2/26/00 1228	3/21/00 0045	-0.1	1.02
ANGI	NO_y	NO	3/21/00 0046	9/2/00 0045	-0.1	1.02
ANGI	NO_y	NO	9/2/00 0046	9/3/00 0045	-0.1	1.03
ANGI	NO_y	NO	9/3/00 0046	9/9/00 0045	-0.1	1.02
ANGI	NO_y	NO	9/9/00 0046	9/10/00 0045	-0.1	1.01
ANGI	NO_y	NO	9/10/00 0046	9/11/00 0045	-0.1	1.02
ANGI	NO_y	NO	9/11/00 0046	9/13/00 0045	-0.1	1.04
ANGI	NO_y	NO	9/13/00 0046	9/15/00 0040	-0.1	1.04
ANGI	NO_y	NO	9/15/00 0041	11/21/00 0045	-0.1	1.08
ANGI	NO_y	NO	11/21/00 0046	2/4/01 0000	-0.1	1.17
ANGI	NO_y	NO_y	9/1/00 0046	9/2/00 0045	-1.18	1.07
ANGI	NO_y	NO_y	9/2/00 0046	9/5/00 0045	-1.18	1.08
ANGI	NO_y	NO_y	9/5/00 0046	9/6/00 0045	-1.18	1.09
ANGI	NO_y	NO_y	9/6/00 0046	9/7/00 0045	-1.18	1.1
ANGI	NO_y	NO_y	9/7/00 0046	9/11/00 0045	-1.18	1.14
ANGI	NO_y	NO_y	9/11/00 0046	9/12/00 0045	-1.18	1.19
ANGI	NO_y	NO_y	9/12/00 0046	9/13/00 0045	-1.18	1.19
ANGI	NO_y	NO_y	9/13/00 0046	9/14/00 0948	-1.18	1.2
ANGI	NO_y	NO_y	9/15/00 0041	11/21/00 0045	-1.18	1.04
ANGI	NO_y	NO_y	11/21/00 0046	2/4/01 0000	-1.18	1.13
ANGI	NO_y	NO_y	2/1/00 0000	2/24/00 1254	-1.18	1.02
ANGI	NO_y	NO_y	2/24/00 1255	2/26/00 1227	-1.18	0.79
ANGI	NO_y	NO_y	2/26/00 1228	3/21/00 0045	-1.18	1.02
ANGI	NO_y	NO_y	3/21/00 0046	8/31/00 0045	-1.18	1
ANGI	NOy	NOy	8/31/00 0046	9/1/00 0045	-1.18	1.06
ANGI	Ozone	O_3	4/1/00 0000	8/24/00 0244	-0.21	1.06
ANGI	Ozone	O_3	8/24/00 0245	8/25/00 0244	-1.02	1.06
ANGI	Ozone	O_3	8/25/00 245	8/29/00 0244	-1.02	1.12
ANGI	Ozone	O_3	8/29/00 245	8/31/00 2359	-1.02	1.21
ANGI	Ozone	O_3	9/7/00 1356	11/26/00 2359	1.28	1.02
ANGI	Ozone	O_3	11/27/00 2359	2/15/00 2359	1.28	1.08
ANG100	NO_y	NO	12/5/00 0000	12/18/00 2359	-0.1	0.99
ANG100	NO_y	NO	12/19/00 0000	2/23/01 1157	0	1.11
ANG100	NO_y	NO_y	12/5/00 0000	12/18/00 2359	-0.7	1

Table 3-7. Calibrations applied to the gaseous pollutant data.

Site	Instrument	Parameter	Start Date/Time PST	End Date/Time PST	Applied Calibration Factors	
					Zero Offset	Slope
ANG100	NO_y	NO_y	12/19/00 0000	2/23/01 1157	-2.6	1.09
ANG100	Ozone	O_3	11/30/00 1139	12/19/00 1044	-0.3	0.99
ANG100	Ozone	O_3	12/19/00 1045	12/22/00 1631	-1.4	0.99
ANG100	Ozone	O_3	12/22/00 1632	1/1/01 1224	0.6	0.99
ANG100	Ozone	O_3	1/11/01 1225	2/15/01 2359	4.5	0.99
BAC	NO_y	NO	2/25/00 1700	3/7/00 0045	-0.688	1.02
BAC	NO_y	NO	3/7/00 0046	4/14/01 0045	-0.02	1.066
BAC	NO_y	NO	4/14/00 0046	5/16/00 1150	-0.02	1.064
BAC	NO_y	NO	5/21/00 0 0046	6/1/00 1155	-0.02	0.902
BAC	NO_y	NO	6/1/00 1156	7/15/00 0045	-0.02	1.091
BAC	NO_y	NO	7/15/00 0046	8/7/00 1200	-0.02	1.19
BAC	NO_y	NO	8/7/00 1300	8/29/00 0045	-0.02	1.136
BAC	NO_y	NO	8/29/00 0046	8/31/00 2359	-0.02	1.091
BAC	NO_y	NO	9/1/00 0000	9/18/00 0045	0.17	1.091
BAC	NO_y	NO	9/18/00 0046	12/7/00 2359	0.17	1.148
BAC	NO_y	NO	12/8/00 0000	2/28/01 2359	-0.05	1.148
BAC	NO_y	NO_y	2/25/00 1700	3/7/00 0045	3.2	0.97
BAC	NO_y	NO_y	3/7/00 0046	4/14/01 0045	-1.02	1.025
BAC	NO_y	NO_y	4/14/00 0046	5/16/00 1150	-1.02	1.065
BAC	NO_y	NO_y	5/21/00 0046	6/1/00 1155	-1.02	0.921
BAC	NO_y	NO_y	6/1/00 1156	7/15/00 0045	-1.02	1.043
BAC	NO_y	NO_y	7/15/00 0046	8/7/00 1200	-1.02	1.12
BAC	NO_y	NO_y	8/7/00 1300	8/29/00 0045	-1.02	1.087
BAC	NO_y	NO_y	8/29/00 0046	8/31/00 2359	-1.02	1.041
BAC	NO_y	NO_y	9/1/00 0000	9/18/00 0045	-1.09	1.041
BAC	NO_y	NO_y	9/18/00 0046	2/28/01 2359	-1.09	1.086
BTI	NO_y	NO	11/18/00 1504	2/15/01 0559	-0.03	1.09
BTI	NO_y	NO_y	11/18/00 1504	2/15/01 0559	-0.77	1.05
SNFH	NO_y	NO	11/24/00 0000	11/30/00 0908	no change	1.18
SNFH	NO_y	NO	11/30/00 1923	12/24/00 1323	no change	1.06
SNFH	NO_y	NO	12/24/00 1405	2/10/01 2359	no change	0.9
SNFH	NO_y	NO_y	11/24/00 0000	11/30/00 0908	-0.17	1.12
SNFH	NO_y	NO_y	11/30/00 1923	12/24/00 1323	-0.17	1.02
SNFH	NO_y	NO_y	12/24/00 1405	2/10/01 2359	-0.17	0.87
SNFH	Ozone	O_3	12/1/00 0000	1/10/01 2359	1.13	1.04
SNFH	Ozone	O_3	1/11/01 0000	1/16/01 2359	1.13	0.98
SNFH	Ozone	O_3	1/17/01 0000	1/19/01 2359	1.13	0.96
SNFH	Ozone	O_3	1/20/01 0000	2/13/01 1441	1.59	0.96
SNFH	HNO_3	HNO_3	12/2/00	12/16/00 0045	0.17	1.00
SNFH	HNO_3	HNO_3	12/16/00 0046	12/28/00 0029	0.17	0.97
SNFH	HNO_3	HNO_3	12/28/00 0030	1/12/01 0045	0.21	0.97

Table 3-7. Calibrations applied to the gaseous pollutant data.

Page 3 of 3

Site	Instrument	Parameter	Start Date/Time PST	End Date/Time PST	Applied Calibration Factors	
					Zero Offset	Slope
SNFH	HNO_3	HNO_3	1/12/01 0046	1/24/01 0045	0.21	0.94
SNFH	HNO_3	HNO_3	1/24/01 0046	1/28/01 0045	0.15	0.94
SNFH	HNO_3	HNO_3	1/28/01 0046	2/15/01 1200	0.15	0.92
BAC	SO_2	SO_2	11/20/00 1934	12/29/00 0920	1.15	1.01
BAC	SO_2	SO_2	12/29/00 0921	1/3/01 1200	0.43	1.01
BAC	SO_2	SO_2	1/3/01 1201	1/5/01 0259	1.15	1.01
BAC	SO_2	SO_2	1/5/01 0300	2/15/01 2359	1.15	1.03

The baseline values for the NO_y-HNO_3 parameter from the Sierra Nevada Foothills site had a diurnal variation that made it necessary to apply offsets that varied linearly between selected daytime periods. A modification was made to the SurfDat program that allowed the user to select two zero offset values, constituting the range (minimum, maximum) of the variation, and the time period over which to apply the linearly extrapolated values. Only at this site and only for the NO_y-HNO_3 parameter did zero offsets vary in this way.

3.5.4 Documenting Changes to the Data

SurfDat gives the option of prompting the user for log comments during editing; for data validation of the CRPAQS data, this option was always chosen so that any and all changes made to the data were documented. The information contained in the SurfDat log file (see Figure 3-11) included the following: (1) the time the QC technician created the comment, (2) site, (3) flag start time, (4) flag end time, (5) action performed, (6) comment, (7) software version, and (8) method of validation – batch or manual. A prompt for a log comment appeared every time changes to data values (i.e., in applying zero and slopes) or QC codes were made. Log files were updated (and appended to) every time the CDF file was saved. If a particular data point was flagged more than once, the comment for the previous flag was not overwritten; this allowed multiple comments to be attached to a single record. Log files were automatically saved in the same folder as the corresponding CDF file to ensure comments were exported back into the SQL database with the data.

3.6 UNIQUE DATA VALIDATION ISSUES

During validation, every instrument had its own procedures and validation anomalies. This section summarizes our findings by parameter.

3.6.1 Ozone

The overall data validation process for ozone was illustrated in Figure 3-1. Ozone concentration data were queried from the SQL 1-minute database by site (Angiola, Angiola

100-m tower, and Sierra Nevada Foothills) and placed into a SurfDat file for validation. The first step was to ensure all the nightly zero-span calibrations were flagged appropriately. The QC code assignment occurred only for times when an OP code was associated with the data record. Due to various instrument start-up issues and errors in coding, some of the nightly zero-spans were not identified by an OP code, and the corresponding data values were not automatically assigned a QC code during import. In addition, data directly following these scheduled calibrations (periods of ambient data recovery) were not assigned OP codes. The ambient recovery periods were defined by the QC technician as periods in which the records still indicated the presence of calibration gas; the data during these periods were invalidated. The amount of recovery time was, on average, about 5 minutes following the last minute of calibration gas sent. For these zero-span-related data points, QC codes were assigned manually in SurfDat.

The next step in the data validation process involved consulting the instrument's off-line summary to determine and mark periods in which the instrument was off-line. The site operator compiled the off-line summary at the end of the study by consulting a variety of resources, including the site and instrument logs. The summary contained information on problem start time, instrument off-line and on-line times, whether the work done was routine or repair, a description of the work performed, the impact of the problem, and a recommended procedure for fixing the data. Flagging these data was accomplished using the SurfDat batch-edit process, including verifying the changes made by examining the time series plots. After the data technician reviewed off-line summary information, any unresolved issues in the data were followed up by referring to the original documentation including ozone instrument logs, instrument task sheets, and site logs. If questions were still unanswered, daily QC comments and e-mails from the field staff were reviewed. The ozone analyzer manual was consulted in cases in which data points were impossible or unlikely to result given the instrument limitations. The measurement expert was able to provide advice regarding the probability that given data were "real." As a last resort, the data technician investigated the raw ozone data files from which the data originated.

For selected periods of time, ozone data and NO/NO_y data were investigated together. A CDF file containing both sets of data were created and brought into SurfDat. "Suspect" time periods when the ozone concentrations dipped were often the result of a natural titration with NO and were, therefore, actually valid. During times of interest or pollution events, data from the ozone analyzer were compared (usually via web site plots) to data from most other instruments at a particular site, in order to determine what data were similarly affected and the nature of the event.

The time scale used in examining the time series plots was usually set to 12 to 48 hours. When a period of interest was identified, the time scale was often narrowed to the hour level. Time series of longer periods (up to 15 days for the ozone data) were normally viewed after validation at the micro-level. When longer term data were reviewed, for example over a period of weeks, trends in the data could be more easily identified, such as instrument drift and calibration responses.

QC codes 0 to 3 and 6 to 9 were used to flag ozone data. Calibrations, zero-spans, filter changes, scheduled work, and maintenance were all flagged with the QC code 2 ("calibration/instrument check"), but data values remained unchanged. Depending on the severity of a known instrument problem, data were flagged with a QC code of 3, indicating "equipment problem" (the data value remaining unchanged) or 8, indicating "invalid" (data value changed to a null code). Although data flagged with a QC code of 2 or 3 were technically "invalid," the use of these flags allowed the record to retain the data value, which could be used in assessing calibrations, evaluating matrix zero air data, or investigating other issues.

QC code 7 ("suspect") was commonly applied to periods of data when events occurring at the site (e.g., maintenance checks, instrument cleaning, etc., as noted in the off-line summary and instrument log) may have affected data but had the appearance of being "real." For odd data periods when no information was available, data were also flagged as "suspect" and accompanied with a comment explaining the reasoning behind the suspicion. If the suspect data point was beyond the scope of the instrument's ability to resolve (e.g., out of range, impossible point-to-point variation, etc.), it was flagged as "invalid" (QC code 8) instead.

A QC code of 8, signifying invalid data, was usually used in cases in which calibrations, work on the instrument, and equipment problems could not explain the invalid nature of the data. Care was taken implementing this flag because the data value was changed to a null value. (Note however, that the original RawValue is always available in the STI database; only the ModValue can be changed).

Table 3-8 provides a summary of validation comments and QC code changes (from most used to least used) for ozone at Angiola.

In SurfDat, all records flagged with a QC code of 8 or 9 had a data value changed to null (-980). The only other changes made to the data value were the result of applying zero offsets and slopes. The zero and span values applied to the ozone data were previously reported in Table 3-7.

3.6.2 NO/NO_y

The validation process for NO/NO_y data was nearly the same as that for ozone data. One-minute NO and NO_y data were exported from the SQL 1-minute database as a SurfDat CDF file. For the first step, nightly zero-spans and matrix zero-air checks (four times daily) were investigated and flagged appropriately when QC code translation failed or was not implemented. The period of ambient data recovery following a calibration or matrix zero air check was also flagged. On average, about 5 minutes of data were invalidated to account for recovery.

Instrument off-line summaries were consulted prior to the examination of site logs, instrument task sheets, and instrument logs. The NO_y analyzer manual, daily QC comments, e-mails from the field staff, and raw NO and NO_y data files were reviewed to resolve any unanswered questions. Discussion with the measurement expert provided insight into any outstanding issues, and all validated data were reviewed and approved by the expert before they were returned to the SQL database.

Table 3-8. Common validation comments and QC code changes for ozone data (example is from Angiola).

Comment	Description
Missing data	QC code set to 9
Nightly zero-span calibration	QC code set to 2
Ambient recovery following nightly zero-span cal	QC code set to 2
Filter change	QC code set to 2
Filter change and recovery	QC code set to 2
Single value excursion >20% of surrounding values and not supported by NO data	QC code set to 8 and data value set to -980
Zero offset; zero=-1.28[a]	Data value was offset by 1.28
Span signal problem	QC code set to 8 and data value set to -980
Slope=1.06	Data value was multiplied by 1.06
Missing data: power failure	QC code set to 9
Slope=1.02[b]	Data value was multiplied by 1.02
Zero	QC code set to 2
Manual zero-span	QC code set to 2
Okay	QC code set to 0
Multipoint calibration	QC code set to 2
Manual calibration	QC code set to 2
O_3 instrument stopped working correctly; flatlined	QC code set to 8 and data value set to -980
Manual zero-span cal; clock reset	QC code set to 2
New scrubber installed	QC code set to 2
Data recovering from power failure	QC code set to 8 and data value set to -980
Instrument problem	QC code set to 3
Instrument being repaired	QC code set to 8 and data value set to -980
Solenoid valve changed out	QC code set to 2
Unknown	QC code set to 8 and data value set to -980
Audit	QC code set to 2
Filter change; valve check; inlet check	QC code set to 2
Calibration with instrument check	QC code set to 2
Ambient recovery following filter change; inlet and valve checks	QC code set to 2
[O_3]=0; odd spike	QC code set to 7
[O_3] dip	QC code set to 7
Calibration/instrument check	QC code set to 2
Slope larger than QC limits	QC code set to 7
Span from 830-1018 (Dennis) to check O_3 output	QC code set to 2

[a] Several other zero offset values were used as well.
[b] Several other slope values were used as well.

Table 3-8. Common validation comments and QC code changes for ozone data (example is from Angiola).

Comment	Description
Switching valve leak; bad scrubber (instrument goes off-line for repairs)	QC code set to 3
Odd [O_3] spike	QC code set to 7
O_3 data affected by NO/NO_y calibration	QC code set to 2
Uncharacteristic [O_3] spike	QC code set to 7
Unknown	QC code set to 7
Missing data; DAS off-line	QC code set to 9
Instrument put back on-line following repairs	QC code set to 3
Instrument problems; work being done on instrument	QC code set to 3
Negative spike; [O_3] drops from 34 ppb (1059) to -8 (1100)	QC code set to 8 and data value set to -980

[a] Several other zero offset values were used as well.
[b] Several other slope values were used as well.

QC codes 0 to 3 and 6 to 9 were used to flag NO and NO_y data. Calibrations, zero-spans, matrix zero-air checks, filter changes, scheduled work, and maintenance were all flagged with a QC code of 2 (“calibration/instrument check”), but data values remained unchanged. “Equipment problem” corresponded to QC code of 3, and periods when known instrument problems occurred were flagged this way rather than as “invalid” so that data values were retained. QC codes 1 (“estimated”) and 6 (“failed QC”) were rarely used with this data set.

QC code 7 (“suspect”) was commonly applied, especially during periods when events were occurring that very well may have affected the data, although data values remained near ambient levels. Odd periods of data, when no information supporting the suspicion was available, were sometimes labeled as suspect. The QC technician’s reasoning for the QC code change was also documented. Suspect data points occurring outside the instrument’s range or data values unlikely to occur in nature were flagged as “invalid” (QC code 8). QC code 8 was applied only in instances in which the QC technician was absolutely certain of (and had relevant information to support) invalidation of the data.

Table 3-9 describes some criteria used in flagging NO/NO_y data.

For the period December 13-17, 2000, the Bethel Island NO/NO_y instrument had numerous operational problems and data were invalidated. For November 18 through December 20, 2000, the nightly calibration system at Bethel Island was not plumbed properly to adequately document instrument performance on a daily basis; data were flagged as suspect.

Table 3-9. Example NO/NO_y QC code changes and comments.

Flag	QC Code	Issue	Example Comment
Suspect	7	NO>NO_y (where NO-NO_y<3 ppb)	[NO]>[NO_y]
Suspect	7	six consecutive equal values	Six consecutive equal values
Suspect	7	NO and/or NO_y exceeds 700 ppb (urban site)	NO and/or NO_y exceeds 700 ppb (urban site)
Suspect	7	NO and/or NO_y exceeds 300 ppb (rural site)	NO and/or NO_y exceeds 300 ppb (rural site)
Suspect	7	Large point to point variation (>30 ppb)	Large point-to-point variation
Invalid	8	Large point to point variation (>>30 ppb)	Large point-to-point variation
Invalid	8	Large drops, spikes	Odd spike following prolonged (failed) matrix zero air
Invalid	2	Matrix zero-air check	Matrix zero-air
Invalid	2	Nightly zero-span calibration	Nightly zero-span calibration
Invalid	2	Ambient recovery following calibration	Ambient recovery following calibration
Invalid	2	Filter change	Filter change
Invalid	2	Biweekly calibration	Biweekly calibration

3.6.3 Nephelometer

Five-minute nephelometer b_{sp} data, along with the nephelometer sampling chamber temperature, RH, and pressure (where available) data, were validated using SurfDat. The data validation criteria applied to the nephelometer data are discussed below and summarized in **Table 3-10**. The QC codes noted in the discussion are the codes stored with the data in the STI database. The translations from the STI QC code to the ARB primary validation code are also provided in Table 3-10.

Nephelometer zeroes, spans, and miscellaneous calibrations were not automatically flagged by the data system. Information provided by the off-line summaries and instrument logs were used along with the time series plot of the data to assess the periods when calibrations were performed. Calibrations were flagged with an STI QC code of 2, which translates to the ARB primary validation code of "I" for invalid.

All b_{sp} readings greater than 2000 Mm^{-1} were invalidated (STI QC code 8), because the electronic circuitry of the instrument was incapable of generating valid data beyond that value. Positive spikes were not invalidated or flagged as suspect unless appropriate information was provided for flagging. Negative spikes, however, were usually invalidated (when additional information was provided) or flagged as suspect (when no information was available). Spikes or troughs in the nephelometer b_{sp} data were compared to collocated BAM $PM_{2.5}$ mass concentrations where available.

Table 3-10. Quality control criteria applied to the nephelometer data.

Phenomena	QC category	STI QC code	ARB QC code
b_{sp} data greater than 2000 Mm^{-1}	Invalid	8	I
b_{sp} data recorded when RH is greater than 75%	Flagged	4	S
Large dips in b_{sp}, temp, pressure, and/or RH	Invalid	8	I
Negative b_{sp} values	Invalid	8	I
Extreme scatter in b_{sp} values	Invalid - equipment problems	3	I
Lack of variation in b_{sp}	Invalid - equipment problems	3	I
Instrument zeros, spans, and routine maintenance	Invalid - routine maintenance	2	I
Data collected at a position or elevation other than what is specified in the site description (affects Angiola Tower data).	Flagged	5	I
Gaps in the 5-minute data filled in with time averaged 1-minute data	Flagged	10	V0
Data collected with heater on continuously	Flagged	11	S

There were a number of periods when the heater for the nephelometer instrument was inoperative or operating incorrectly. When RH values rose over 75%, an STI QC code of 4 was assigned to flag the data. In other instances, the heater stayed on continuously. These data were given an STI QC code of 11. Both STI QC codes 4 and 11 translates to the ARB primary validation code of 'S' for suspect.

The DAS software for daily ingestion of the digital nephelometer data (the GetNeph program) often locked up during its scheduled run near midnight every evening. Occasionally, a downward spike would be present in the data at the exact minute when the program ran. These data spikes were invalidated. Another indication of instrument lock-up or GetNeph program failure was a constant, unchanging b_{sp} value over time. Data from these "flatline" periods were invalidated.

The 5-minute digital data of the Radiance Research nephelometer, acquired by the GetNeph program, were used as the primary data source. The analog signals from the nephelometer were collected by the CRPAQS DAS on a 1-minute time basis as a backup to the digital DAS. Thus, 1-minute analog data may exist for periods when the digital data were missing. Data gaps in the digital nephelometer data in excess of 12 hours were filled in using the analog data whenever possible. Filled data were given an STI QC code of 10, which translates to the ARB primary validation code of 'V0' for valid. The 1-minute analog data corresponding to a period to be filled were reviewed in SurfDat using the same procedures as the 5-minute digital nephelometer data. Five-minute averages were created from the validated 1-minute data using a 75% data completeness criterion. Only valid 1-minute data were included in the 5-minute average. The analog data compare well with the digital data from the nephelometer with no apparent biases. Thus, the averaged analog data were used directly to fill in missing values over the selected periods. Times when 1-minute analog data were used to fill in gaps in the 5-minute digital data are listed in **Table 3-11.**

Table 3-11. Periods when the 5-minute serial nephelometer data were filled with 5-minute averages of the 1-minute analog data.

Site	Date/Time Start	Date/Time End
Angiola	1/22/00	2/3/00
Angiola	5/8/00 1450	5/9/00 1455
Angiola	5/22/00 0840	5/26/00 0005
Walnut Grove	12/14/00 0005	12/22/00 1250
Sacramento Del Paso	1/27/00 2355	2/1/00 1155
Sacramento Del Paso	2/1/00 1430	2/3/00 1740
Sacramento Del Paso	5/19/00 2350	5/21/00 0005
Sacramento Del Paso	7/21/00 1855	8/11/00 1455
Sacramento Del Paso	11/17/00 1305	11/24/00 8050
Sacramento Del Paso	12/6/00 1210	12/7/00 5050
Sacramento Del Paso	1/11/01 1555	1/12/01 9050
Sacramento Del Paso	2/5/01 1600	2/9/01 2355

The Angiola Tower nephelometers at 50 m and 100 m were lowered to ground level for maintenance activities. Data within the time periods when the nephelometers were not at their standard positions were flagged with the STI QC code of 5. These data are considered invalid and have the ARB primary validation code of "I".

The Angiola trailer and the Angiola Tower 1-m nephelometers were close enough to be considered collocated. The b_{sp} values from these two nephelometers consistently compared well. **Figure 3-13** shows a scatter plot of the Angiola Tower 1-m versus the Angiola trailer nephelometer during the entire period the 1-m nephelometer was operational.
Angiola 1-m b_{sp} data collected between December 22, 2000, and January 2, 2001, were flagged as suspect after comparison with the trailer nephelometer.

The four Angiola nephelometers were essentially collocated during periods when the tower platforms were lowered to the ground for maintenance. The valid b_{sp} values from the 100-m, 50-m-, and trailer nephelometers were compared during multiple periods when the tower was down. These comparisons yield inconsistent results. **Figures 3-14** shows a time series plot of the four Angiola nephelometers b_{sp} data for February 5 through February 9, 2001. This period, when the tower instruments were at ground level, was chosen because it is relatively prolonged, and overnight readings would be free of any site-operator effects on the data. The trailer and the 1-m nephelometers are highly correlated during this period, with a reasonable intercept of 1.5 Mm^{-1}, a slope of 0.88, and an R^2 of 0.99 (see **Figure 3-15)**. The cause of bias between the 1-m and Angiola trailer nephelometers is yet to be determined. The trailer and 100-m nephelometers compared poorly over this period with a slope of 0.44, an intercept of 10.2

Mm^{-1}, and an R^2 of 0.74 (see **Figure 3-16)**. The 50-m nephelometer also compared poorly with the trailer nephelometer, shown in **Figure 3-17**, with an unrealistic intercept of -7.0 Mm^{-1}, a slope of 0.74, and an R^2 of 0.87. The nephelometers involved in the February 2001 comparisons were the same used in the March 2001 inter-comparison study that produced excellent results (Richards et al., 2001a; 2001b).

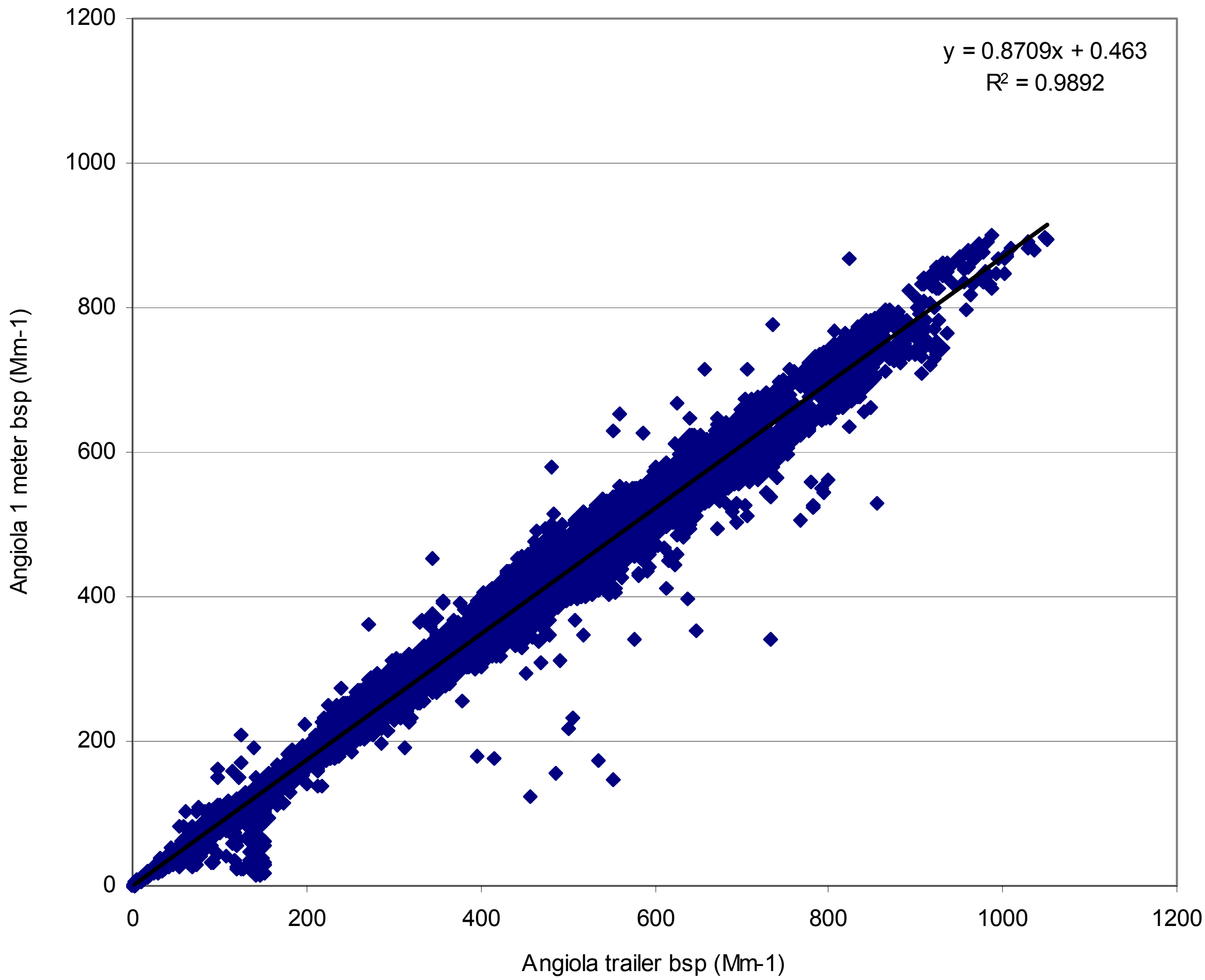


Figure 3-13. Comparison of the Angiola 1-meter nephelometer and the trailer nephelometer b_{sp} (Mm^{-1}) for December 14, 2000 through February 9, 2001. Only valid data are shown.

We have no explanation as to why the 100-m and 50-m nephelometers do not compare well with the 1-m and trailer nephelometers. We flagged as suspect all the data collected by both the 100-m and 50-m nephelometers. Although the data from any of the tower nephelometers produce b_{sp} data that are suitable for the evaluation of general conditions at the Angiola site, the suitability of tower data to evaluate differences in the atmospheric conditions due to elevation needs to be assessed through further analysis of the data.

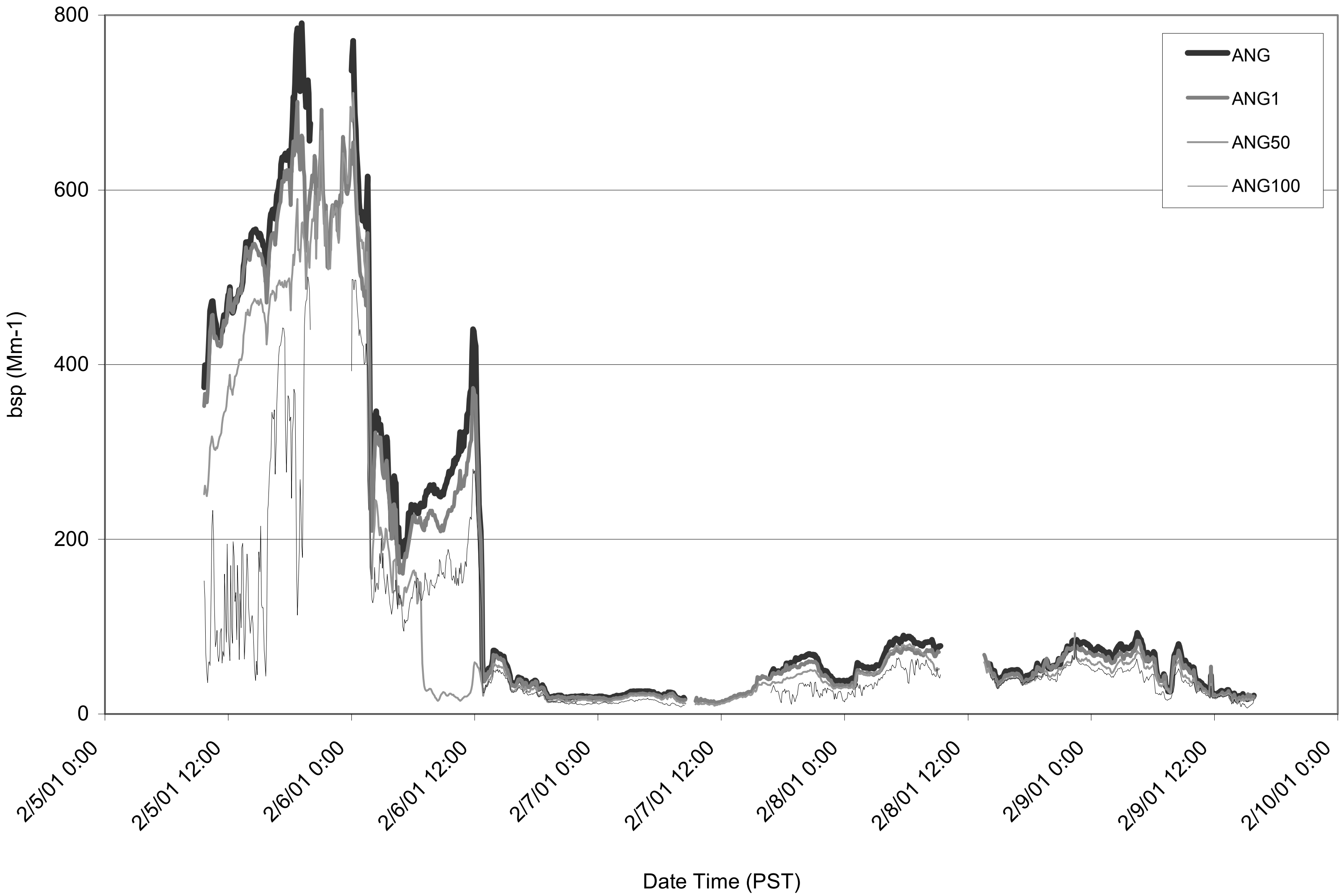


Figure 3-14. Time series plot of b_{sp} (Mm^{-1}) from the Angiola trailer nephelometer (ANGI) and the Angiola Tower nephelometers at 1 m, 5 m, and 100 m, (ANG1, ANG50, and ANG100, respectively). All data collected during times of instrument calibration and maintenance have been removed.

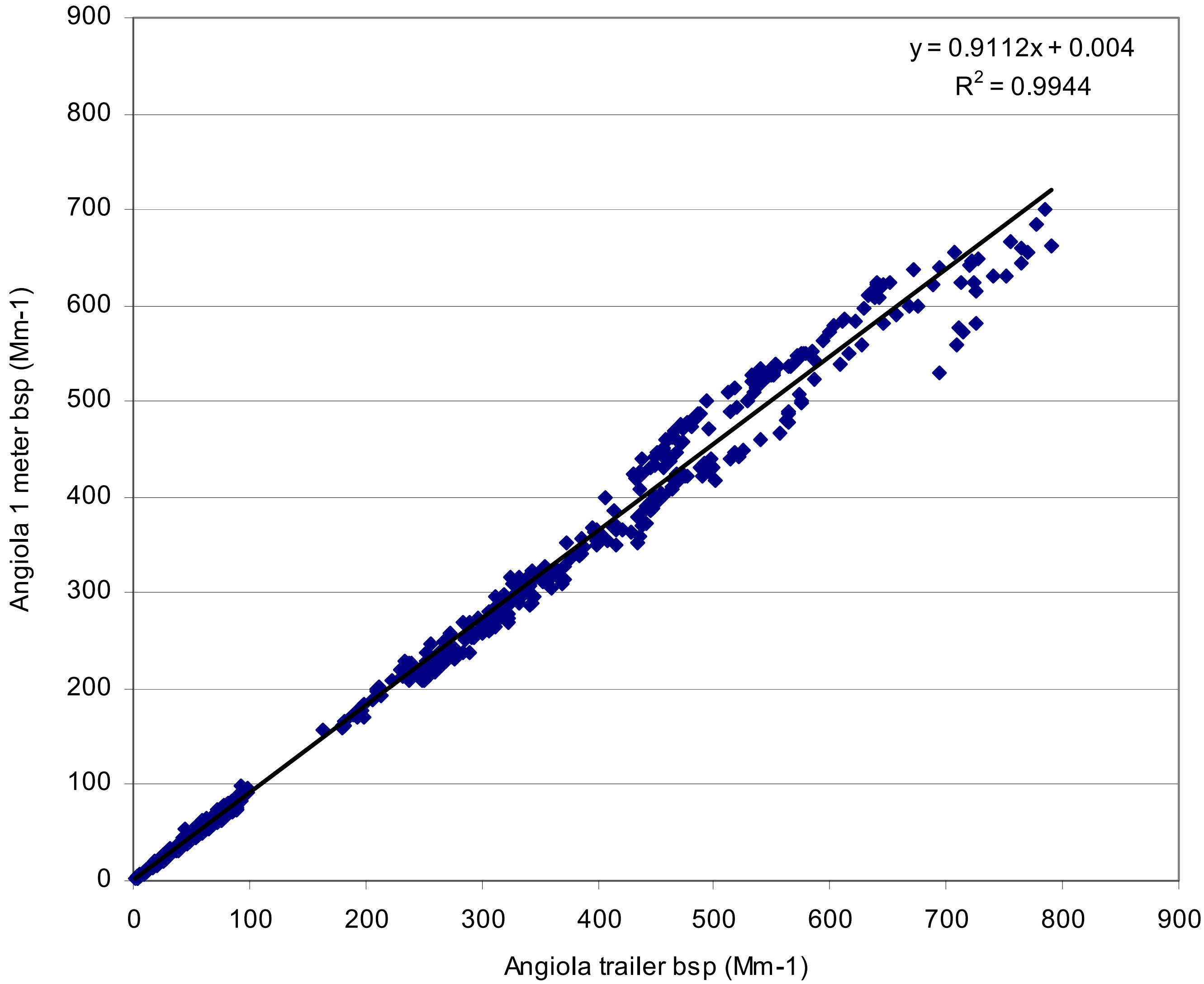


Figure 3-15. Comparison of the Angiola 1-m nephelometer and the trailer nephelometer b_{sp} (Mm^{-1}) for February 5 through February 9, 2001. All data collected during times of instrument calibration and maintenance have been removed.

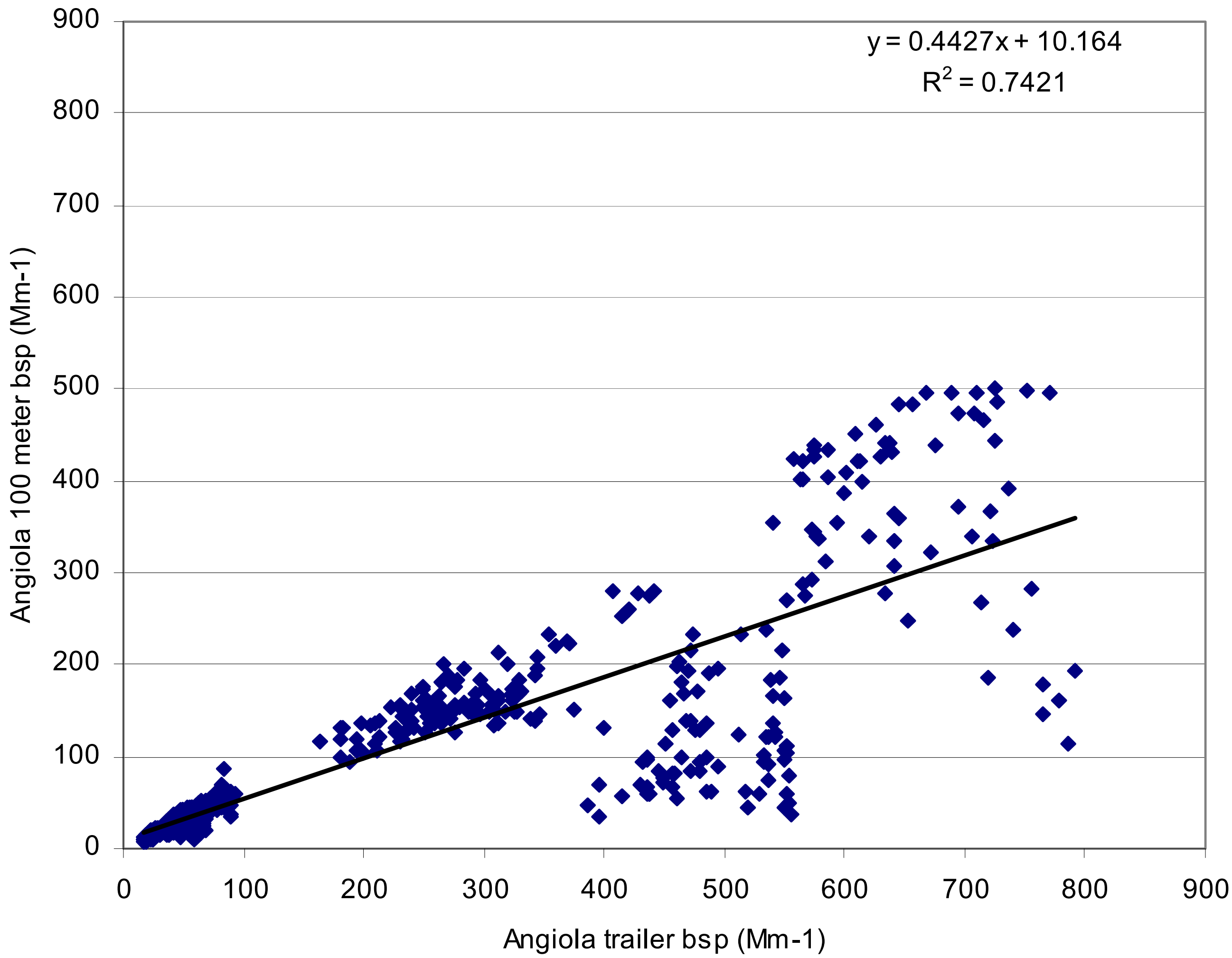


Figure 3-16. Comparison of the Angiola 100-m nephelometer and the trailer nephelometer b_{sp} (Mm^{-1}) for February 5 through February 9, 2001. All data collected during times of instrument calibration and maintenance have been removed.

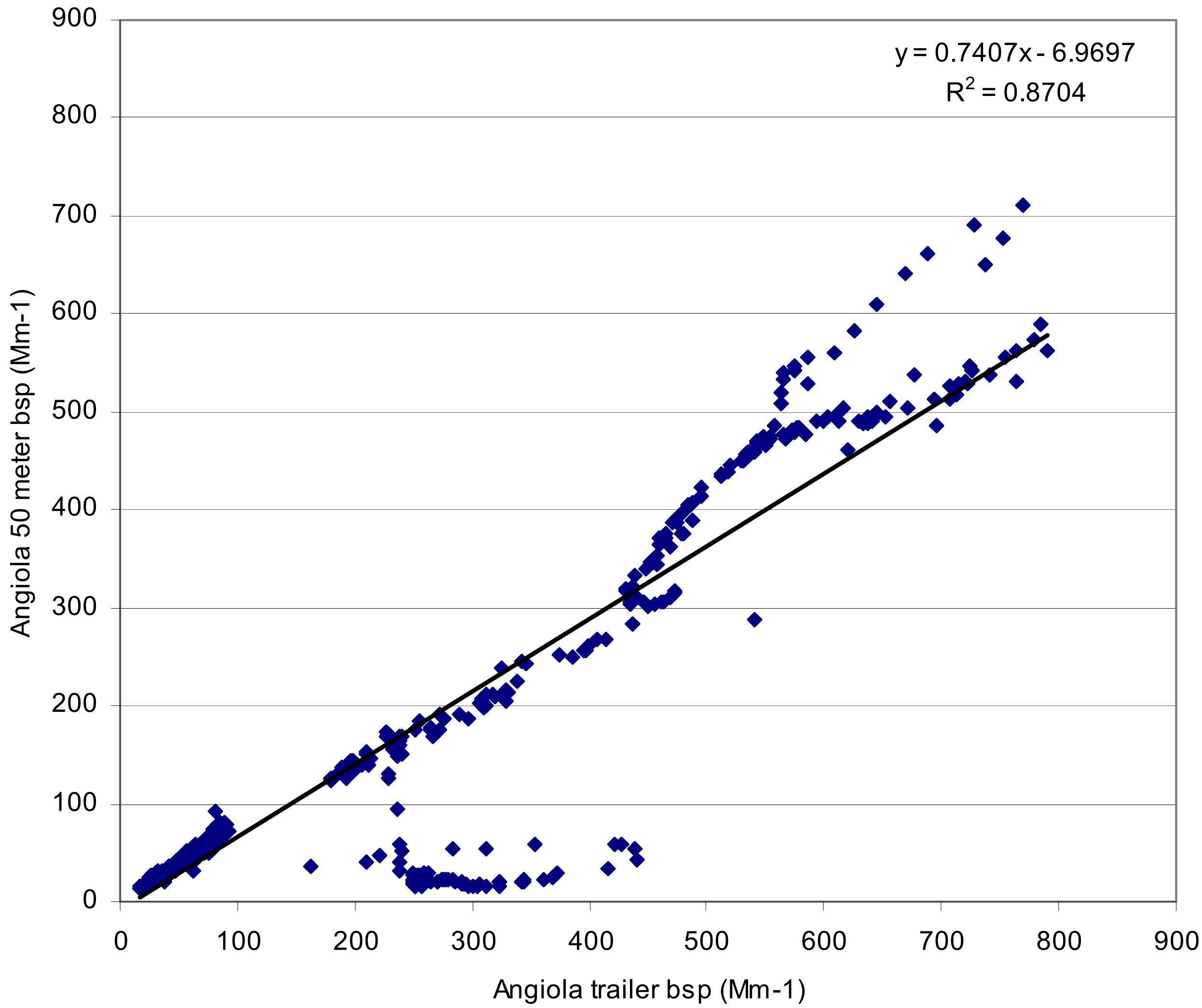


Figure 3-17. Comparison of the Angiola 50-m nephelometer and the trailer nephelometer b_{sp} (Mm^{-1}) for February 5 through February 9, 2001. All data collected during times of instrument calibration and maintenance have been removed.

3.6.4 BAM PM_{10} and $PM_{2.5}$

PM_{10} and $PM_{2.5}$ concentration sample volume data were extracted from the 60-minute database by site and brought into SurfDat for validation[3]. All sites except Sacramento and San Jose 4th Street sampled at 16.67 standard liters per minute (SLPM) and reported concentrations in standard conditions (µg/m^3 at 273.15 K and 760 mm Hg). Sacramento and San Jose instruments both sampled volumetrically at 16.67 LPM_{actual} and reported data corrected to ambient conditions (µg/$m^3{}_{actual}$).

[3] Volume data were used for validation purposes only.

Off-line summaries, BAM field site logs, instrument logs, and instrument task sheets were reviewed to identify when calibrations were performed, general maintenance tasks were done, or instrument problems existed. Concentrations and flow volume data were flagged for all off-line dates and times when equipment problems, instrument maintenance, flow calibrations or audits were noted. Volume and concentration data were also invalidated for severe equipment problems (i.e., tape transport errors, pump failure, power failure) which were documented in site or instrument logs.

All field calibration data from monthly site flow and leak checks were compiled into tables to track the operation of the instruments' flow meters and plumbing. These summaries included the results from the independent performance audits. We used the following screening criteria:

- When the difference between the flow rate measured with the certified flow meter and with the instrument was greater than 10% ($\pm$10% of 835 L, or less than 751.5 L or greater than 918.5 L), the concentration data were flagged as suspect.
- When the difference between the flow rate measured with the certified flow meter and the instrument was greater than 20% ($\pm$20% of 835 L, or less than 668 L, or greater than 1002 L), the concentration data were flagged as invalid.

PM_{10} and $PM_{2.5}$ concentration data were viewed together for sites at which both parameters were measured. During periods of interest or when unusual data were observed, collocated Aethalometer BC data and nephelometer b_{sp} data were compared to the BAM concentrations. Additional screening criteria included the following:

- The data were invalidated when flow volume = 0 L and mass concentration = 0 $\mu g/m^3$.
- All negative concentration values were flagged as suspect.
- Three or more consecutive identical concentration values were flagged as suspect.
- Instances in which $PM_{2.5}$ concentrations exceeded PM_{10} concentrations were flagged as suspect.
- Mass concentration data were invalidated for large point-to-point variations when other site data did not support the large differences (i.e., Aethalometer or nephelometer).
- Instances of mass concentration equal to 995 $\mu g/m^3$ were flagged as suspect because this is the maximum range of the instrument. The instrument manufacturer suggests that this high concentration may indicate condensation on the measurement tape but 995 $\mu g/m^3$ values were observed in the summer when condensation was unlikely in the SJV.

After inspecting and flagging data in SurfDat, all the data were reviewed by the instrument expert prior to placement back into the 60-minute database.

Monthly calibrations, and at least one audit during the study, at all sites produced data that were flagged with an STI QC Code of 2 which translates to the ARB primary validation code of "I" for invalid. Typical reasons for flagging concentration data as suspect, STI QC code 7, were negative concentration values, $PM_{2.5}$ concentration greater than PM_{10} concentration values, and more than three consecutive identical values. STI QC code 7 translates to the ARB primary validation code of “S” for suspect. Instrument problems were site specific and are discussed next. Data collected during events such as tape transport or power failures were flagged with STI QC code 3 which translates to the ARB primary validation code of “I” for invalid.

There were several instrument changes during the study; these changes are documented in **Table 3-12**.

Table 3-12. Summary of BAM instrument changes by quarter.

SiteId	Equipment Description	S/N	4th Qtr 1999	1st Qtr 2000	2nd Qtr 2000	3rd Qtr 2000	4th Qtr 2000	1st Qtr 2000	On-line Date/Time	Off-line Date/Time
ALT1	Model 1020 BAM - PM25	x4946		x	x	x	x		1/28/00 1800	12/8/00 1100
ALT1	Model 1020 BAM - PM25	y2133					x	x	12/8/00 1600	2/8/01 0800
ANGI	Model 1020 BAM - PM10	x4152	x	x					1/20/00 8000	3/7/00 1200
ANGI	Model 1020 BAM - PM10	x4619		x	x	x	x	x	3/7/00 1200	1/19/01 1900
ANGI	Model 1020 BAM - PM10	x4619						x	1/23/01 1400	2/6/01 1000
ANGI	Model 1020 BAM - PM25	x4619	x	x				x	1/20/00 0700	3/7/00 1200
ANGI	Model 1020 BAM - PM25	x4152		x	x	x	x	x	3/7/00 1300	1/19/01 1800
ANGI	Model 1020 BAM - PM25	x4619						x	1/19/01 1900	1/23/01 1300
ANGI	Model 1020 BAM - PM25	x4152						x	1/23/01 1300	2/6/01 0900
BAC	Model 1020 BAM - PM10	x4153		x	x	x	x	x	1/21/00 2300	2/6/01 1500
BAC	Model 1020 BAM - PM25	x5217		x	x	x	x	x	1/21/00 2300	2/6/01 1500
BTI	Model 1020 BAM - PM25	y2133					x		11/17/00 1300	12/8/00 1100
BTI	Model 1020 BAM - PM25	x4946					x	x	12/8/00 1200	2/15/01 400
COP	Model 1020 BAM - PM10	y2134				x	x		9/13/00 1500	11/14/00 2200
COP	Model 1020 BAM - PM25	y2133				x	x		9/13/00 1400	11/14/00 2200
EDW	Model 1020 BAM - PM10	y2134			x	x			6/20/00 1900	9/1/00 0700
EDW	Model 1020 BAM - PM25	y2133			x	x			6/20/00 1900	9/1/00 0700
SDP	Model 1020 BAM - PM25	y1484			x	x	x	x	4/13/00 1500	2/7/01 0800
SJ4	Model 1020 BAM - PM25	y1722			x	x	x	x	5/18/00 1800	2/15/01 2200
SNFH	Model 1020 BAM - PM25	y2134					x	x	11/19/00 1300	2/12/01 1200

Altamont Pass (ALT1)

The Altamont Pass $PM_{2.5}$ BAM experienced nearly monthly tape transport errors. One instrument change occurred during the study.

Angiola (ANGI)

BAM instruments were switched twice during the study at Angiola. Several problems occurred at the Angiola site that affected data validation. The dominant data validation issue was the frequent occurrence of power failures during the winter months. Differences between actual and measured flow rates also occurred at Angiola:

- The difference between the actual and measured flow rates for the BAM PM_{10} instrument on July 24, 2000 was greater than 10%. Volume and concentration data from July 19, on which date a flow calibration was performed, through July 24, 2000, when the flow rate difference was noted, were flagged as suspect.
- The difference between the actual and measured flow rates for the BAM $PM_{2.5}$ instrument on February 22, 2000, was greater than 10%. Volume and concentration data from January 28, on which date a flow calibration was performed, through February 22, 2000, were flagged as suspect.

In both of these cases, the indicated flow rate may have been more than 10% higher than the true flow rate due to a flow meter problem.

During the months of January and February 2000, the $PM_{2.5}$ concentrations exceeded PM_{10} concentrations on many occasions; these data were flagged as suspect.

Finally, to assist data analysts, numerous periods during which high PM_{10} concentrations occurred were noted. A few data points were flagged as suspect because of large point-to-point variations or because the concentrations reached the maximum range of the instrument.

Bethel Island (BTI)

There was one instrument change during the study at Bethel Island (Table 3-12). The Bethel Island BAM $PM_{2.5}$ instrument failed a flow calibration audit on December 5, 2000, due to a leak in the system. Consequently, flow volume and concentration data from the previous flow calibration on November 18 to December 5, 2000, were flagged as suspect.

Corcoran (COP)

The Corcoran BAM $PM_{2.5}$ instrument experienced a tape transport error for several days from September 28 to October 3, 2000; these data were all flagged as invalid.

Sacramento (SDP)

The Sacramento BAM $PM_{2.5}$ instrument experienced numerous power failures in May and June 2000 (resulting in missing data) and some tape transport errors in June, October, and December 2000 (resulting in invalid data).

The difference between the actual and measured flow rates for the BAM $PM_{2.5}$ instrument on December 8, 2000, was greater than 10% due to a temperature probe problem. The indicated flow rate may have been up to 10% higher than the true flow rate. Volume and concentration data from the previous flow calibration on July 14 through December 8, 2000, were flagged as suspect.

Sierra Nevada Foothills (SNFH)

The Sierra Nevada Foothills BAM instrument had a data acquisition problem from December 19 to December 22, 2000, at which time the program chip was replaced. Data from this time period were flagged as invalid.

San Jose (SJ4)

The San Jose BAM instruments experienced intermittent power failures in May 2000. In addition, the difference between the actual and measured flow rates for the BAM $PM_{2.5}$ instrument on July 17, 2000 was greater than 10% due to a flow meter problem. The indicated concentration values may have been up to 13% higher than the true flow rate. Volume and concentration data from the previous flow calibration on June 23 through July 17, 2000, were flagged as suspect.

Other Sites

There were no major issues with the BAM data at Edwards Air Force Base or Bakersfield. Of interest to analysts is the observation that numerous periods with high PM_{10} concentrations were observed at Bakersfield.

3.6.5 Aethalometer

Black carbon data were obtained from two models of the Aethalometer, a 1-wavelength instrument and a 7-wavelength instrument. The data validation approach was similar for both instrument types and is summarized in this section. Aethalometer BC and flow data were pulled from the 1-minute SQL database. Nephelometer b_{sp} data were also used (as QC checks) in the validation process.

Off-line summaries, Aethalometer field site logs, instrument logs, and instrument task sheets were reviewed to compile a summary of flow check results (Section 2.5.5). Periods exceeding 12 hours of missing data that were noted in off-line summaries and/or instrument logbooks were documented (**Tables 3-13 and 3-14**). The data were briefly reviewed in SurfDat to verify this information and to identify undocumented periods of missing data. In addition, the following factors were documented: major repairs or modifications to the instrument that would affect data (collected before or after repairs/modifications); sampling start and end dates; instrument serial number(s); and instrument exchanges (when instrument substitutions were made).

Table 3-13. Periods of missing 1-wavelength Aethalometer data greater than 12 hours.

Site Code	Start of Missing Data	End of Missing Data	Reason
ANGI	1/16/00 0000	1/16/00 2355	Unknown
ANGI	2/28/00 1645	2/29/00 0800	Unknown
ANGI	9/24/00 2323	9/26/00 0823	Unknown
BODB	1/23/01 0645	1/25/01 1500	Power failure and DAS did not reboot properly
BODB	2/10/01 1850	2/14/01 1422	Power failure and DAS did not reboot properly
SDP	5/06/00 2340	5/11/00 1205	Power failure and DAS did not reboot properly
SDP	5/12/00 2355	5/13/00 0959	Power failure and DAS did not reboot properly
SDP	5/15/00 1645	5/17/00 1255	DAS failure
SDP	5/21/00 0105	5/22/00 1055	DAS failure
SDP	5/25/00 0105	5/25/00 1150	DAS failure
SDP	5/30/00 0105	5/30/00 0915	DAS failure
SDP	6/02/00 0105	6/02/00 1125	DAS failure
SDP	6/06/00 2355	6/07/00 1120	DAS failure, DAS repaired
SJ4	8/01/00 1425	8/10/00 0902	Unknown
SJ4	8/11/00 1025	8/17/00 1614	Vacuum hose disconnected
WAG	11/22/00 0000	11/22/00 2355	Unknown

The following steps were taken to validate the Aethalometer data:

- Checked all parameters for values that were a common fraction (e.g., ½) of the expected value. If these data followed a pattern (e.g., flow rate value immediately proceeding a tape advance is 1.7 LPM), the parameter was adjusted and the change was documented.
- Completed the set of QC checks listed in **Table 3-15** by referring to site/instrument logbook, task sheets, and off-line summaries to determine the appropriate periods in which to invalidate the data.
- Reviewed data visually in SurfDat for high concentration spikes.
  - **One-wavelength data**: When abnormally high concentrations occurred in isolation, the BC concentrations were checked for correlation with nephelometer b_{sp} data. If the nephelometer data agreed, the concentration spike was documented in the QC summary. If nephelometer data did not agree, the spike was documented in the QC summary and a comment was included about the disparity.

Table 3-14. Periods of missing 7-wavelength Aethalometer data greater than 12 hours.

Site Code	Start of Missing Data	End of Missing Data	Reason
ANG 100	none >12 hours		Two power failures 1/1/01 1220-1815; 1/8/01 0705-1110
ANGI	12/26/2000 1013	12/27/2000 1428	Unknown
ANGI	1/19/2001 1708		7-wave fails, 1-wave (S/N 229) installed
ANGI		1/30/2001 1136	7-wave reinstalled
ANGI	2/5/2001 0555	2/7/2001 0000	Unknown
COP	11/8/2000 1614	11/9/2000 0900	Circuit breaker tripped - repaired and on-line after 0900, 11/09/00; in off-line summary
EDW	7/14/2000 1800	7/21/2000 1919	Unknown
SDP	10/23/2000 0031	10/31/2000 0 840	Tape transport error -- repaired and on-line 10/31/00 1150
SDP	11/16/2000 2155	11/20/2000 1425	Instrument had restart problems.
SDP	11/26/2000 1955	11/27/2000 1205	Unknown
SDP	11/28/2000 2355	11/30/2000	Restart problems – listed in log as 11/27/00 to 11/30/00 1520
SDP	12/22/2000 1800	12/25/2000 1220	Restart problem - in log, unknown start time through 12/25/00 1215
SDP	12/28/2000 0110	12/29/2000 1655	Tape transport error and restart problems
SDP	12/30/2000 1625	1/3/2001 1310	Aeth failed to reset. Installed relay to perform reset and activated 1/3/01 1246.
SDP	1/3/2001 1930	1/5/2001 1545	Aeth restarting. Installed Power booster and advanced tape manually.
SDP	1/7/2001 0025	1/8/2001 0930	Unspecified instrument problem
SJ4	10/22/2000 0700	10/24/2000 0902	Unknown
SJ4	11/1/2000 2109	11/2/2000 0855	Stop/error light on aeth from unknown problem; restarted ok
SJ4	11/15/2000 0649	11/18/2000 1443	Unknown problem; operator replaced tape and diskette
SJ4	1/17/2001 1144	1/18/2001 1102	Unknown
SJ4	1/21/2001 2020	1/22/2001 1111	Unknown
SNFH	12/4/2000 1408	12/6/2000 0916	Removed instrument S/N 271; replaced with S/N 272
SNFH	12/20/2000 2159	12/21/2000 0944	Tape advance error - tape came off of take-up spool after checks 12/20/00

Table 3-15. Validation criteria and associated action for Aethalometer data.

Parameters to Flag in a Record	Reason for Flagging	QC Code	QC Definition	Comments
All wavelengths and flow rate	Instrument problems (irregular)	3	Invalid	Repairs, power failure, flow leak, etc.
All wavelengths and flow rate	Routine maintenance	2	Invalid	Flow check, optical test strip, etc.
All wavelengths and flow rate	Flows <0.01 SLPM	2	Invalid	Tape advance
All wavelengths and flow rate	Flow problems	3	Invalid	Flow >±20% of 6.9 SLPM (flow < 5.52, flow >8.28)
All wavelengths and flow rate	Flow problems	7	Suspect	Flow >±10% of 6.9 SLPM (flow < 6.21, flow > 7.59)
All wavelengths, individually	Negative concentration <-0.5 μg/m^3	3	Invalid	Equipment problem
All wavelengths, individually	Slightly negative concentration <-0.1 μg/m^3	7	Suspect	Possible equipment problem; instrument noise

 – **Seven-wavelength data**: When abnormally high concentrations occurred in isolation, the wavelength concentrations were checked for correlation with one another. If all wavelength concentrations tracked together, contemporaneous nephelometer b_{sp} data were checked. If nephelometer data agreed, the spike was documented in the QC summary. If nephelometer data did not agree, the spike was documented in the QC summary and a comment was included about the disparity. If concentrations at different wavelengths did not correlate (e.g., concentrations at one wavelength spiked while other wavelength concentrations remained low), a comment was included in the QC summary, and the data point was flagged as suspect (QC code 7) because of aberrant concentration(s).

- Checked the data for large (>0.5 SLPM) point-to-point variations in the flow rate. Checked also for data variations that accompanied small flow variations in an exaggerated manner (e.g., flow increased by 0.1 LPM and concentration increased by 0.5 μg/m^3).
- Reviewed data visually in SurfDat for abnormally low concentrations. Checked for low concentrations that hovered near zero and/or did not track with concentrations from other wavelengths immediately following tape advances. These very low, aberrant concentrations sometimes persisted for a few minutes to a few hours after a tape advance. These data were documented in the QC summary and flagged with an invalid QC code of 3. If there was a "transition" value (following clearly aberrant data and not yet quite in line with other wavelengths or an expected pattern), this value was flagged as suspect (QC code 7).
- Checked the data for high concentrations that did not track with other wavelengths immediately following tape advances. The QC code was not changed, but the instruments and wavelengths in which this problem existed were documented.

After inspecting and flagging data in SurfDat, all the data were reviewed by the instrument expert prior to placement back into the 1-minute database.

Site-specific Aethalometer issues are discussed below.

Angiola (ANGI)

One-wavelength: Throughout the duration of the site operation, several periods of invalid data occurred because of insects found in the optical chamber. From the time that an insect entered the instrument until the time that site operators removed it, data were invalidated. The time that the insects entered the optical chamber of the instrument was obvious from the erratic spikes, made up of both positive and negative concentrations, of the data values. In March and April 2000, the instrument inlets were covered with screen material to eliminate this problem.

Seven-wavelength: The flow pump began failing at the beginning of the Fall study and was rebuilt by November 9, 2000. Before the rebuild, the flow rate gradually dropped; data records were flagged as suspect or invalid when flow volume fell below 10% or 20%, respectively. The instrument software was upgraded on January 17, 2001. All data became erratic on January 18, 2001; the following day, the instrument was taken off-line. In its place, a 1-wavelength Aethalometer was installed and came on-line on January 20, 2001. The 7-wavelength Aethalometer was reinstalled on January 30, 2001.

Angiola Tower (ANG100)

Seven-wavelength: Whereas most wavelengths seldom showed significant negative concentration values, the 590-nm wavelength had 127 values between -0.5 and -0.1 $\mu g/m^3$. From February 10, 2001, at 1832 to February 12, 2001, 0302, the concentrations measured at the 590-nm wavelength vacillated around the concentrations measured at other wavelengths, always staying within ~0.5 $\mu g/m^3$, but often dipping below zero (the BC concentrations were low during this period). We have no explanation for this behavior and data were not flagged.

A variation in concentration at this site appeared to be directly related to the flow. On occasion, when the flow increased slightly (e.g., 0.1 SLPM), the concentration also increased, with the increase in concentration more noticeable and of larger magnitude than the flow change. A small decrease in flow did not have an analogous affect.

Long periods of data were invalidated because the tower instrument carriages were lowered to the ground for maintenance. These periods ranged from a few hours to a few days in length and are summarized in **Table 3-16**.

Bakersfield (BAC)

One-wavelength: As with the Angiola site, several periods of data during February and March 2000 were invalid due to instrument error and insects found in the optical chamber. Additionally, during periods in July 2000, the flow rates dipped below 90% of the set point flow, and the data were flagged as suspect.

Table 3-16. Times that the Angiola tower instrument carriages were lowered to the ground (Tower Down) and returned (Tower Back Up); pertinent to ANG100 data only.

Tower Down	Tower Back Up
8/21/2000 1135	8/21/2000 1515
8/31/2000 1235	8/31/2000 1535
9/12/2000 0745	9/12/2000 1430
10/5/2000 1035	10/5/2000 1330
10/10/2000 0800	10/12/2000 1200
10/17/2000 0810	10/17/2000 1125
10/19/2000 0940	10/19/2000 1335
10/23/2000 1305	10/24/2000 1145
10/31/2000 0945	10/31/2000 1350
11/7/2000 1000	11/7/2000 1700
11/16/2000 1140	11/22/2000 1755
11/28/2000 1520	12/1/2000 1800
12/4/2000 1015	12/5/2000 1925
12/6/2000 1315	12/8/2000 1945
12/13/2000 0910	12/14/2000 1520
12/19/2000 0940	12/23/2000 1945
12/29/2000 1205	12/29/2000 1835
1/10/2001 1530	1/11/2001 1935
1/18/2001 0930	1/19/2001 1950
1/22/2001 0920	1/22/2001 1010
1/26/2001 0935	1/26/2001 1600
1/30/2001 0950	1/30/2001 1845
2/5/2001 0935	2/9/2001 1550

Seven-wavelength: When Aethalometer 256 (serial number) was at BAC, concentrations at 520-nm wavelength consistently read about one-third lower than all other wavelengths. A similar trend was found at EDW using the same instrument. A lamp alignment/intensity problem could be responsible for this disparity. Concentrations measured at the 520-nm wavelength on instruments 254 and 255 displayed a similar trend.

Flow audits performed on February 6 and February 14, 2001, indicated that flow rates were 32.5% to 41% low. One task sheet indicates that the flow was adjusted on February 6, but it does not indicate the magnitude of the adjustment. The most recent prior flow check was performed on January 29, 2001. All data between the January 29, 2001 flow check and February 14, 2001, have been marked with a QC code of 7.

Bodega Bay (BODB)

One-wavelength: At the Bodega Bay site, the power was intermittent during the end of January and beginning of February 2001. The DAS did not reboot correctly after two power failures, and no data were available for January 23-25 and February 10-14, 2001.

Bethel Island (BTI)

Seven-wavelength: The concentrations measured at 660- and 950-nm wavelengths periodically showed high-frequency (low amplitude) vacillations, as well as erratic high concentrations following tape advances. The 520-nm wavelength consistently read 20% to 30% lower than all other wavelengths; this effect was particularly pronounced at high concentrations.

The initial flow audits on January 3, 2001, indicated that instrument flow was more than 10% high. Though the flow meter was then recalibrated and another flow audit performed, all data from the previous flow audit on December 12, 2000, until January 3, 2001, were flagged as suspect.

The data from the 520- and 590-nm wavelengths were highly erratic during January 8-11, 2001. During this time period, two power outages occurred. All values from the 520-nm wavelength during this period were flagged as suspect.

Corcoran Patterson Avenue (COP)

Seven-wavelength: The data collected at the 590- nm wavelength of instrument 271, specifically when it was at the Corcoran site, vacillated much more markedly than the data from other wavelengths. For example, during certain sampling periods, every fourth value was slightly negative. While at any given time the concentrations measured at the 590-nm wavelength did not differ greatly from the mean concentrations at the other wavelengths, the 590-nm concentrations fluctuated rapidly and regularly.

Sacramento Del Paso (SDP)

One-wavelength: Several operational issues complicated data collection and retrieval at this site. Throughout the months of May and June 2000, the DAS failed regularly until it was replaced on June 7, 2000. During periods when the DAS failed, there were no data available until site personnel went to the site and rebooted the system. Additionally, flows that were less than 90% of the target flow rate during the first months of site operation, from January through March 2000, resulted in many periods of suspect data.

Seven-wavelength: Following tape advances, the 470-nm wavelength concentration frequently did not track with other wavelength concentrations. After a tape advance, the 470-nm concentration hovered about zero for a few minutes to a few hours. Eventually its concentration jumped up to correlate with the other wavelength concentrations. The concentrations measured at the 590-nm wavelength vacillated and were abnormally high following tape advances. This problem was unique to instrument 271 and a similar pattern was found when this instrument was at SNFH.

<u>**San Jose (SJ4)**</u>

One-wavelength. At the San Jose site, the Aethalometer was changed out several times because of instrument problems. In April 2000, a discrepancy of at least 10% was found between instrument-indicated flow and the volumetric flow rate, as measured by the BIOS, the volumetric flow standard. It is not known if there was a sudden change in flow rate or whether the flow measurement slowly degraded over the previous month. Data from this period of low flow were validated according to the flow standard results. The instrument was later replaced.

<u>**Sierra Nevada Foothills (SNFH)**</u>

Seven-wavelength: The concentrations in the 590-nm wavelength of instrument 271 (which was at SNFH through December 4, 2000) fluctuated consistently and rapidly. Some highly negative values of approximately -30 $\mu g/m^3$ were observed; these values were invalidated. With instrument 271, the 470- and 660-nm wavelength concentrations frequently did not track with other wavelength concentrations after tape advances. During these periods, the 470- and 660-nm concentrations hovered around zero for a few minutes to a few hours. Eventually, though not necessarily simultaneously, their concentrations jumped up and correlated with other wavelength concentrations.

<u>**Walnut Grove (WAG)**</u>

One-wavelength: At the Walnut Grove site, a bias in the reported data following tape advances was observed. For 10 to 20 minutes after a tape advance, reported data values increased. The data values during this period were only marginally higher than neighboring data. It is possible that the reported data values were 10% to 20% higher than actual values, as discussed in Section 2.5.5. Although it is possible that the bias occurred throughout the site operation, this trend was only observable throughout the month of November and the beginning of December 2000 at this site.

3.6.6 OCEC

OC and OCEC concentration and instrument volume data were pulled from the 60-minute database by site and brought into SurfDat for validation. Prior to validation, all data, except missing data, had a QC code of zero (0). As validation progressed, validation actions were limited to data that had QC codes of zero, in order to avoid overwriting other non-zero QC codes. The parameters reported to ARB comprised OC mass concentration and OCEC mass concentration.

Off-line summaries, OCEC field site logs, instrument logs, and instrument task sheets were reviewed to compile a summary of flow-check results. Periods exceeding 12 hours of missing (or invalid) data that were noted in off-line summaries and/or instrument logbooks were documented and are summarized in **Table 3-17**. The data were briefly reviewed in SurfDat to verify this information and to identify undocumented periods of missing data. In addition, the

Table 3-17. Periods of at least 12 consecutive hours of missing OCEC data. Note that no valid data were collected at Angiola from 4/18/2000 1000 to 11/03/2000 1200.

Site	Start Date/Time (inclusive)	End Date/Time (inclusive)
BAC	2/15/2001 2100	2/16/2001 2100
ANGI	2/28/2000 1500	2/29/2000 6000
ANGI	2/6/2001 1200	2/7/2001 8000

following factors were documented: major repairs or modifications to the instrument that would affect data (collected before or after repairs/modification including modification of dwell times or temperature plateaus); sampling start and end dates; instrument serial number(s) and instrument exchanges (when instrument substitutions were made); and dwell times and burn temperature plateaus (see Table 2-11).

The following steps were taken to validate the OCEC data:

- Completed the set of QC checks listed in **Table 3-18**. Referred to site/instrument logbook, task sheets, and off-line summary to determine appropriate periods to flag data as suspect or invalid.
- Investigated large point-to-point sample volume variations. One-time differences of more than 50 L (5% of expected value) were examined and flagged as suspect; however, these large jumps usually signaled other significant instrument problems.
- Perused afterburner temperatures for consistency. Theoretically, the temperature should have been consistently ~750°C, but the value was almost always 50°C. It was suspected that the actual afterburner values did not transmit accurately from the instrument. Inconsistent afterburner temperature values often coincided with abnormal instrument operation; thus, consistency was more important in validation than were specific values.
- Scanned data for multiple consecutive identical points. If more than three consecutive identical values were present, the identical values were flagged as suspect.
- Investigated data for unrealistic high values. Values over 20 $\mu g/m^3$ were individually evaluated and compared with contemporaneous BAM and Aethalometer data.

Neither flow nor concentration calibration factors were applied to the data. If calibration results did not fall within the acceptable range, affected and potentially affected data were invalidated. After inspecting and flagging data in SurfDat, all the data were reviewed by the instrument expert prior to placement back into the 60-minute database. Summaries of site-specific data validation issues follow.

Table 3-18. Validation criteria for OCEC.

Parameters to Flag in a Record	Grounds	QC Code	QC Definition	Comments
OC, OCEC	Documented instrument problems	3	Invalid	Repairs, power failure, etc.
OC, OCEC	Routine maintenance	2	Invalid	Flow audit, CO_2 audit, etc.
OC, OCEC	OP code > 0	3 or 7	Invalid or suspect	Internal indication of instrument problem
OC, OCEC	Volume <0.01 L	3	Invalid	No sample flow or reporting error
OC, OCEC	Volume >±20% 1000 L (V > 1200; V < 800)	3	Invalid	Flow path obstruction, operator interference, instrument error
OC, OCEC	Volume >±10% 1000 L (V > 1100; V < 900)	7	Suspect	Flow path obstruction, operator interference, instrument error
OC, OCEC	OC>OCEC by more than 0.1 $\mu g/m^3$	3	Invalid	Software glitch or other instrument problem
OC, OCEC	OC>OCEC by less than 0.1 $\mu g/m^3$	7	Suspect	Cycle aborted or mass concentration below detection limit
OC, OCEC	OC < -0.1 OCEC < -0.1	3	Invalid	Equipment problem
OC, OCEC	OC < 0.001 OCEC < 0.001	7	Suspect	Possible meteorological effects; unrealistic data

Bakersfield (BAC)

Sample volumes were reported not transmitted to the DAS from startup on October 4 through 1100 PST on November 3, 2000, and from 1200 PST on November 10 to 1800 PST on November 14, 2000. However, there were no log notes, task sheets, or notes in the off-line summary that indicated any intervention or action taken on the instrument on either November 3 or 10, 2000. Sample volumes during these times were checked on the backup CD, and data were validated accordingly.

Throughout the instrument's time at Bakersfield, burn temperatures were recorded as 275°C and 750°C for the plateaus P3 and PF, respectively. Burn (dwell) times[4] D3 and DF were recorded as 300 and 360 seconds, respectively (see Table 2-11). These settings were not consistent with the Angiola settings.

Data from startup on October 4 through 1200 PST on October 20, 2000, were invalidated because sample volumes from instrument start-up were very erratic, frequently dipping well

[4] Duration of burn plateau

below 800 L. Also prior to October 20, 2000, the D3 was decreased from 480 to 300 seconds on instructions from the measurement expert.

In our validation of the data, we noted that EC, not OC, seemed to drive the carbon spikes at Bakersfield. Contemporaneous Aethalometer data affirmed the spikes and often agreed with the EC value to within ±1 $\mu g/m^3$.

Angiola (ANGI)

Sometime prior to February 23, 2000, instrument S/N 5400B202579908 was at this site; however, this instrument had multiple problems, and no valid data were collected from it. By February 23, 2000, a new instrument (S/N 5400B202699912) was installed. Carbon dioxide audits were partially performed and had reasonable results; however, no initial flow audit was documented. A flow audit was not performed until April 12, 2000, and the instrument passed this audit. Thus, valid data potentially began as early as February 25, 2000 (after some initial installation checks). Table 2-11 lists operating parameters during this time; they differ somewhat from operating parameters at other times and sites in the study.

Because of the different operating parameters and lack of documentation, all data from 2200 PST on February 25 through 1000 PST on April 18, 2000, were flagged as suspect. Data from 1000 PST on April 18 through 2000 PST on November 3, 2000, were invalidated. Reasons include incorrect settings, abnormal settings, presumed afterburner heating problems, thermocouple failure and replacement, suspiciously low data (OC concentrations were 0.00 $\mu g/m^3$ for many consecutive records), failed zero air calibrations, burner replacement, skipped cycles, non-zero OP codes, disconnected inlets, software installations, incorrect flow audit calculations, failed leak checks, and pump failure. In addition, there was a dearth of documentation of both routine maintenance and corrective action. As one example, on April 19, 2000, the afterburner temperature was changed from 750°C, the default value, to 0°C; the documentation proffers no explanation for this change.

On October 27, 2000, the instrument was documented as on-line and functional. However, no data came through the DAS until October 30, 2000. There were no notes in the logbook on October 30, 2000, to explain the missing data. As of October 27, flows were slightly low (A = 16.0 LPM, B = 15.8 LPM) though technically in acceptable range.

A flow audit was recorded to have taken place on November 28, 2000, from 0950 to 1330 PST, but no supporting documentation exists. Another flow check, purportedly performed on December 22, 2000, also lacked supporting documentation.

No data collected after the official end of the study (February 6, 2001) were validated.

3.6.7 SO_2

One-minute SO_2 concentration data from the Bakersfield site were brought into SurfDat. Nightly zero-spans were examined and flagged when OP code-to-QC code translation failed. Ambient recovery data, lasting for a period of about 5 minutes following calibrations, were also flagged appropriately. The remainder of the validation process closely followed that of ozone and NO/NO_y data reduction. SO_2 data were investigated together with these other gaseous species, as well as with a variety of other parameters during times of interest (e.g., suspected pollution events). Data were viewed on a 12- to 48-hr time scale, reducing the scale to specific data points and enlarging it to view a longer time scale when necessary.

At Bakersfield, data from November 28 to December 4, 2000 (0559 PST), exhibited excessive noise (2 to 3 ppb) near baseline levels; these data were flagged as suspect.

3.6.8 PAN/NO_2

For NO_2, data were flagged as suspect if calibration factors varied by more than 20% per day. Outliers were identified by visual inspection of the data. For nearly all the sites, more than half the data were either suspect, invalid, or missing because of instrument problems.

For PAN, the raw 1-minute chromatograms were averaged over 15-minute intervals. The data were visually inspected to determine the magnitude of the PAN peak height above background. Manual peak height determinations were required because there was only a small difference between the PAN concentration and baseline noise (thus making automated background subtraction unreliable). Only data for the intensive operating periods (IOP) of December 26-28, 2000 and January 4-7, 2001 were validated by CE-CERT.

3.6.9 Nitric Acid

NO_y and NO_y-HNO_3 data were queried from the 1-minute database. The first round of data reduction was performed on these parameters in SurfDat, using information derived from off-line summaries, instrument logs, task sheets, etc. Zero offsets and slopes were applied to the data based on nightly zero-span and calibration information. The NO_y and NO_y-HNO_3 data were returned to the SQL database, and nitric acid was subsequently calculated from the difference (between the NO_y and NO_y-HNO_3) and appended to the 1-minute database. The calculated nitric acid and NO_y were pulled together into a new SurfDat file for additional validation.

NO_y data from the NO_y instrument and NO_y data from the nitric acid instrument were compared to one another to see how well the two instruments tracked. Nitric acid and NO_y data from the nitric acid instrument was also compared with other gaseous parameters, such as ozone.

Nitric acid concentrations are derived from the difference between NO_y concentrations and NO_y with nitric acid removed (NO_y minus HNO_3). The nitric acid concentrations were very low compared to the NO_y concentrations. Thus, nitric acid concentrations were subject to considerable uncertainty. For example, peak nitric acid concentrations on the order of 1 ppb might be expected; and, with NO_y concentrations typically 25 to 50 ppb during high days, the nitric acid concentrations would only be 2% to 4% of the NO_y concentrations. The two channels of the instrument would have to be perfectly matched in time to detect the nitric acid.

The results showed that the NO_y-HNO_3 channel of the instrument sporadically gave concentrations of a few ppb when exposed to zero air (either synthetic or matrix) whereas the NO_y channel appeared to respond properly. Possible causes may be that

- The NaCl-coated denuder removes nitric acid by forming $NaNO_3$. High humidity conditions may cause this salt to migrate through the denuder and break off at the outlet and be swept into the analyzer. This is consistent with the high zero air concentration found primarily in the morning hours.
- A localized source of a strong acid such as HCl may also volatilize collected $NaNO_3$.
- The denuder may retain nitrogenous species that slowly volatilize from the denuder.

The combination of low nitric acid concentrations relative to NO_y and the erratic performance of the NO_y-HNO_3 channel resulted in unreliable nitric acid concentration measurements from both nitric acid instruments (Angiola and Sierra Nevada Foothills). The lack of expected diurnal concentrations in nitric acid also supports this conclusion. Based on these observations, all nitric acid data were flagged as invalid. Data from the NO_y channel of the instruments were retained.

From January 20 through January 26, 2001, at Angiola, the operation of the nitric acid instrument monitor and/or calibration system was erratic. Various logbook entries indicate instrument problems but do not clearly identify corrective action. The resulting NO_y data were flagged as suspect.

Converter efficiencies for the NO_y-HNO_3 converter were 100% during the sampling program.

3.6.10 Nitrate

Aerosol Dynamics Inc. (ADI) performed the data validation of the R&P 8400N continuous nitrate data. Data were corrected, flags were applied, and standard deviations were calculated for each point. ADI reprocessed the data from the original instrument files because additional instrument readings were needed to correct the data that were not originally stored by the DAS. The start times used came directly from the instrument. The instrument clock and the site DAS clocks agreed to within four minutes. Sequential corrections were applied to the data as discussed below. Each step is described separately and identified as NO3_1 through NO3_6. Note that all nitrate concentrations are in units of $\mu g/m^3$.

<u>Raw Recalculation (NO3_1)</u>

Due to a calculation error in the instrument software, raw nitrate values were recalculated from the integrated NO_x signals during the flash periods (flash areas) according to Equation 3-1 for 100% theoretical recovery.

$$NO3\_1 = ((FlsArea - BslnArea) * 60)/(ThConvFact * t\_smp * Qsmp) \qquad \textbf{(3-1)}$$

where

FlsArea = Flash Area (ppb*s), the integrated NO_x signal during the flash period.
BslnArea = Baseline Area (ppb*s), the integrated NO_x signal immediately preceding the flash.
ThConvFact = Theoretical Conversion Factor (ppb*s/ng), calculated by the instrument software and confirmed by ADI.

$$\textbf{ThConvFact} = \frac{\text{Standard gas volume}\left(\frac{L}{mol}\right)}{\text{Molar Mass NO}_3\left(\frac{g}{mol}\right) * \text{Analyzer Flow}\left(\frac{L}{\min}\right) * \left(\frac{1\min}{60s}\right)} = \left(\frac{s}{g}\right) = \left(\frac{ppb * s}{ng}\right)$$

t_smp = Sample Time (s), duration of sample collection period.
Qsmp = Sample Flow (L/min), sample flow rate.

Aqueous Standards Calibration (NO3_2)

Equation (3-2) was used to correct for aqueous standard calibration:

$$\text{NO3\_2} = \text{NO3\_1} * (100/\text{AqStd}) \quad \textbf{(3-2)}$$

where

AqStd = The average of all the aqueous standard calibrations at each site (as percent recovery), adjusted for Rcell and NO_x analyzer span drift.

NO_x Span Drift Correction (NO3_3)

Equation 3-3 was used to correct for NO_x span concentration drift:

$$\text{NO3\_3} = \text{NO3\_2} * (\text{CalConc}/\text{Audit\_Fill}) \quad \textbf{(3-3)}$$

where

Audit_Fill = The result of a linear-in-time interpolation of these data.
CalConc = The NO span gas concentration at each site.

CalConc and Audit_Fill are in units of concentration (ppb). NO_x analyzer span checks were corrected for Rcell (see Figure 2-22) pressure variation as discussed below. Remaining span check data variation was attributed to NO_x analyzer span drift.

Rcell Correction (NO3_4)

Equation 3-4 was used to correct for differences in the collection/vaporization cell pressure:

$$NO3\_2 = NO3\_1 * (1 + (0.13 * (Prcell\_n - 5.0))) \quad \textbf{(3-4)}$$

where

Prcell_n = the measured Rcell pressure in inches of mercury (in. Hg). Data were adjusted to a reference Rcell pressure of 5.0 in. Hg.

Instrument response was found to decrease by 13% for every 1 in. Hg the Rcell pressure increased. This relationship was determined empirically using data from several sites, including Angiola ground and tower, Bethel Island, and Sierra Nevada Foothills. Typical Rcell pressures were 5 to 6 in. Hg.

Blank Subtraction (NO3_5)

Equation 3-5 was used to correct for blank values

$$NO3\_5 = NO3\_4 - \min[(0.7133(\mu g/m^3) + .05590*NO3\_4), (0.5*NO3\_4)] \quad \textbf{(3-5)}$$

Field blanks were compared to nitrate concentrations at all sites to yield a linear fit of the form Blank($\mu g/m^3$)=0.7133($\mu g/m^3$)+0.05590*NO3_4($\mu g/m^3$), where NO3_4 is the corrected nitrate value. Both blank values and nitrate concentrations were corrected according to the above procedure. This form avoids meaningless negative numbers for values below the detection limit. If the calculated blank was larger than half of the nitrate value, the blank was set to half of the nitrate value (note that the "min" function in Equation 5 takes the smaller of the two values separated by the comma). Field blanks were higher during periods of high nitrate concentration.

Missing and Invalid Values (NO3_6)

Missing values were assigned a value of -999. Invalid values were assigned a value of -980.

Flags

Data were first examined for flash duration changes, and the data were manually flagged as suspect when the flash duration dropped more than 20% from a stable value. Primary and secondary flags were applied to the data by ADI. Comments were included for flagged data. **Tables 3-19 and 3-20** list the flags used by ADI for the nitrate data. These flags were reported to ARB with the data.

Table 3-19. Primary flags applied by ADI to continuous nitrate and sulfate data.

Abbreviation	Explanation
M	Missing
I	Invalid
12	Invalidated by site operator
V6	Valid value but qualified due to non-standard sampling conditions
V1	Valid value but comprised wholly or partially of below-MDL data
V0	Valid value

Table 3-20. Secondary flags applied by ADI to continuous nitrate and sulfate data.

Secondary Flag[a]	Meaning	Corresponding Primary Flag	Comments
ABT	Aborted cycle	I	The cycle was aborted before completion.
PCA	Previous cycle aborted	I	The previous cycle was aborted resulting in contamination of the strip. This can be viewed as a cleaning cycle.
PWR	Power failure	I	A power failure occurred during the cycle.
DFB	Dynamic field blank	I	A filter was placed on the inlet to measure response to particle-free air.
NRE	Non-representative sample	I	This sample was not deemed representative.
DUR	Flash duration	S	Flash duration dropped by more than 20% from stable value.
AWN	Analyzer warning		This flag was applied when the fault light on the front panel of the pulse analyzer was blinking.
TCH	High cell compartment temperature	V6	Sample temperature (T_{box})> T_{amb} + 10 C resulted in possible volatilization loss of particulate nitrate or sulfate. Non-standard sampling conditions
BID	Below instrument detection limit	V1	Analyte produced an instrument response but the reported value was below the calculated instrument detection limit. Validity of reported value may have been compromised.
TXX	Ambient temperature probe failure		Ambient temperature sensor was not installed correctly or indicated an invalid value (normally due to failure in wet weather). For ambient temperature, this corresponds to a reading of less than -40 C.
TNO	Ambient temperature probe not installed		The cycle was run without the ambient temperature sensor installed.

[a] Listed from highest to lowest priority

Table 3-20. Secondary flags applied by ADI to continuous nitrate and sulfate data.

Secondary Flag[a]	Meaning	Corresponding Primary Flag	Comments
FLS	Flash failure		A failure occurred during the flashing of the nichrome flash strip.
SFO	Sample flow out of range		The sample/bypass flow mass flow meter has an operating range of 1-5V. When the meter recorded a reading of less than 0.5V, this error code was reported. Qsmp is the sample flow controlled by a critical orifice and measured by a mass flow meter.
CFC	Cross flow control fail		The cross flow mass flow controller was not installed correctly or was not maintaining its set point flow percentage through the cross and/or orifice flow lines. The cross flow, Qxflo, is normally set to 85% of the analyzer flow and is that portion of the analyzer flow which flows laterally across the strip.
APO	absolute pressure out of range		The absolute pressure sensor was not installed correctly or indicated an invalid value. For absolute pressure, this corresponds to a reading of less than 0.2 atm or greater than 1.2 atm.
APC	Absolute pressure control fail		The absolute pressure controller was not installed correctly or was not maintaining its set point flow percentage through the cross and/or orifice flow lines.
SMP	Sample pressure out of range		The measured sample pressure in the sample flow lines was out of range. This error is flagged if the sample pressure is greater than the ambient pressure divided by 2.
CDP	Cell dP out of range		The measured cell assembly differential pressure was out of range. This error can only occur during the analysis steps. The error is flagged if the cell assembly differential pressure is greater than zero.
NIE	No issues encountered		

[a] Listed from highest to lowest priority

Summary of Validation Results

Data validation results are summarized as follows:

- The 8400N instrument performs a periodic span check to measure the response of the pulse analyzer to a span gas of known concentration. The frequency of span checks is set by the user. For CRPAQS, the span checks were generally performed every second day at 0100 PST. The SOP settings for the span check require the check to last 10 minutes, resulting in the loss of one data point. In some cases, the span check was programmed to require more time resulting in the loss of two data points (two ten-minute cycles).

- Cold weather led to large fluctuations in pump performance at Sierra Nevada Foothills, as indicated by the R-cell pressure on the 8400N instrument. A temporary box enclosure was installed to alleviate the problem. Smaller fluctuations in pump performance were observed at the Angiola trailer and Bethel Island where pumps were also situated outside. Again, covering the pump reduced R-cell variation. R-cell pressures were stable for pumps situated indoors. Data losses greater than 12 hours are summarized in **Table 3-21**.
- The Bethel Island 8400N instrument experienced a flash failure due to a faulty board on January 11, 2001. Attempts to replace the board on January 16 were unsuccessful due to an electrical short. Finally, pulse generator #107 was replaced by #124 on January 18, 2001 and the instrument operated normally.

3.6.11 Sulfate

ADI performed data validation of the R&P 8400S continuous sulfate data. Data were corrected, flags were applied, and one-sigma uncertainties were calculated for each point. The start times used came directly from the R&P 8400S. A comparison was made between the R&P 8400S clock and the site DAS clock. The DAS clock recorded the time the record was sent which was consistently 10 minutes after the sample start time. Once adjusted for this offset, the

Table 3-21. Periods of at least 12 consecutive hours of missing nitrate data.

Page 1 of 2

Site	Start Time (PST)	End Time (PST)	Cause
ANGI	12/23/2000 19:30	12/24/2000 16:30	Flash Strip Failure
ANGI	1/1/2001 3:50	1/2/2001 11:20	Flash Strip Failure
ANGI	1/16/2001 9:10	1/17/2001 14:30	Strip installed incorrectly
ANGI	2/5/2001 0:00	2/5/2001 23:50	Missing Daily File, error in transmission from the DAS
ANGI	2/26/2001 0:00	2/27/2001 23:50	Missing Daily File, error in transmission from the DAS
ANGT	2/4/2001 23:50	2/5/2001 23:40	Missing Daily File, error in transmission from the DAS
ANGT	2/26/2001 0:00	2/27/2001 23:50	Missing Daily File, error in transmission from the DAS
BAC	11/21/2000 12:20	11/28/2000 10:00	Data cable disconnected from the DAS
BAC	2/15/2001 0:10	2/15/2001 23:50	Missing Daily File, error in transmission from the DAS
BTI	12/3/2000 22:00	12/4/2000 14:20	Ambient Temperature Probe Failure
BTI	12/9/2000 21:50	12/10/2000 17:50	Flash Strip Failure
BTI	12/15/2000 21:10	12/16/2000 15:50	Flash Strip Failure
BTI	12/24/2000 16:40	12/25/2000 15:50	Flash Strip Failure
BTI	1/11/2001 9:00	1/18/2001 16:40	Flash Strip Failure, change from Pulse Generator #107 to #124
BTI	1/23/2001 14:30	1/24/2001 12:30	Flash Strip Failure
COP	10/6/2000 12:10	10/9/2000 0:00	Error in data transmission from 8400N to DAS
COP	10/9/2000 15:30	10/11/2000 7:20	Error in data transmission from 8400N to DAS
COP	10/19/2000 23:40	10/22/2000 23:50	Missing Daily Files, error in transmission from the DAS
COP	11/7/2000 23:40	11/9/2000 9:20	Flash Strip Failure

Table 3-21. Periods of at least 12 consecutive hours of missing nitrate data.

Page 2 of 2

Site	Start Time (PST)	End Time (PST)	Cause
FSF1	12/3/2000 6:30	12/3/2000 6:30	Ambient Temperature Probe Failure
FSF1	12/3/2000 6:30	12/7/2000 9:20	Ambient Temperature Probe Failure
FSF1	1/23/2001 18:00	1/25/2001 11:30	Power Failure
FSF2	2/21/2001 23:50	2/22/2001 23:40	Missing Daily File, error in transmission from the DAS
FSF2	3/13/2001 23:50	3/14/2001 23:40	Missing Daily File, error in transmission from the DAS
FSF2	3/16/2001 23:50	3/17/2001 23:40	Missing Daily File, error in transmission from the DAS
SJ4	11/18/2000 19:00	11/28/2000 13:50	Flash Strip Failure
SJ4	12/4/2000 6:50	12/4/2000 6:50	Flash Strip Failure
SJ4	1/3/2001 3:30	1/4/2001 14:30	Flash Strip Failure
SJ4	1/16/2001 23:50	1/19/2001 14:50	Flash Strip Failure
SJ4	2/9/2001 23:50	2/13/2001 9:30	Power Failure
SNFH	11/26/2000 20:50	11/28/2000 23:10	Ambient Temperature Probe Failure
SNFH	2/3/2001 16:50	2/4/2001 13:20	Flash Strip Failure
SNFH	2/6/2001 4:20	2/8/2001 16:00	Flash Strip Failure
WAG	11/17/2000 2:00	11/26/2000 14:40	Flash Strip Failure
WAG	11/26/2000 15:10	11/28/2000 15:00	Miscellaneous Maintenance
WGT	11/21/2000 23:50	11/26/2000 17:30	Maintenance and Installation
WGT	11/27/2000 15:00	12/2/2000 15:10	Maintenance and Installation
WGT	1/19/2001 6:10	1/21/2001 14:10	Flash Strip Failure
WGT	2/4/2001 1:00	2/4/2001 15:10	Flash Strip Failure

DAS and R&P 8400S times agreed to within 4 minutes. Five sequential corrections were applied to the data as discussed below (SO4_1 through SO4_5). All sulfate units are $\mu g/m^3$.

Raw Recalculation (SO4_1)

Due to a calculation error in the R&P 8400S software, raw sulfate values were recalculated from the Flash Areas according to Equation 3-7 for 100% theoretical recovery

$$SO4\_1 = ((FlsArea - BslnArea) * 60)/(ThConvFact * t\_smp * Qsmp) \qquad \textbf{(3-7)}$$

where

FlsArea = Flash Area (ppb*s), the integrated SO_2 signal during the flash period.
BslnArea = Baseline Area (ppb*s), the integrated SO_2 signal immediately preceding the flash.
ThConvFact = Theoretical Conversion Factor (ppb*s/ng), calculated by the R&P 8400S software and confirmed by ADI (see Equation 3-8)

$$\textbf{ThConvFact} = \frac{\text{Standard gas volume}\left(\frac{L}{mol}\right)}{\text{Molar Mass SO}_4\left(\frac{g}{mol}\right) * \text{Analyzer Flow}\left(\frac{L}{\min}\right) * \left(\frac{1\min}{60s}\right)} = \left(\frac{s}{g}\right) = \left(\frac{ppb * s}{ng}\right) \quad \textbf{(3-8)}$$

where

t_smp = Sample Time (s), duration of sample collection period.
Qsmp = Sample Flow (LPM), sample flow rate.

Aqueous Standards Calibration (SO4_2)

Equation 3-9 was used to correct for aqueous standards calibration:

SO4_2 = SO4_1 * (100/*AqStd*) **(3-9)**

AqStd is normally the average of all the aqueous standard calibrations at each site expressed as percent recovery and adjusted for SO_2 analyzer span drift. In this case, because the instrument was still in development and aqueous standard calibration results were questionable, *AqStd* was set to 100. Thus, sulfate concentrations are reported assuming 100% theoretical recovery. The true theoretical recovery remains unclear for this early version of the R&P 8400S.

SO_2 Span Drift Correction (SO4_3)

Equation 3-10 was used to correct for SO_2 span concentration drift:

SO4_3 = SO4_2 * (*CalConc*/Audit_Fill) **(3-10)**

where

Audit_Fill = the result of a linear-in-time interpolation of these data.
CalConc = the SO_2 span gas concentration at each site

CalConc and Audit_Fill are in units of ppb. SO_2 analyzer audit variation was attributed to SO2 analyzer span drift. .

Blank Subtraction (SO4_4)

A comparison was made of field blank concentration versus ambient sulfate concentration over all the sites to yield a linear fit of the form Blank(μg/m^3) =0.41(μg/m^3). If the calculated blank is larger than half of the sulfate value, the blank is set to half of the sulfate value. Equation 3-11 was used to subtract the blank values from the data where the "min" function takes the smaller of the two values separated by a comma in the equation.

SO4_4 = SO4_3 - min[(0.41(μg/m^3), (0.5*SO4_3)] **(3-11)**

Flags

Data were first examined for flash duration changes, and the data manually flagged as suspect when the flash duration dropped more than 20 % from a stable value. Primary and secondary flags were applied to the data by ADI. Comments were included for flagged data. Data were flagged as “I2” when invalidated by the site operator. “I” and NRE” were manually applied for non-representative samples. Tables 3-19 and 3-20 list the flags used by ADI for the sulfate data and reported to ARB.

Summary of Validation Results

Data validation results are summarized as follows:

- The 8400S was still undergoing development during the CRPAQS study. The first commercial unit was installed at Angiola on January 11, 2001 and the sixth commercial unit was installed at Bakersfield on January 12, 2001. Therefore, only a month of sulfate data was collected at these sites.
- Similar to the 8400N instrument, the 8400S instrument performed a span check of the pulse analyzer every second day at 0100 PST resulting in the loss of one or two 10-minute average data points. Data losses greater than 12 hours are summarized in **Table 3-22**.
- Angiola sulfate data were all deemed invalid because of lack of confidence in the particular instrument installed at the site: 1) this was the first commercial 8400S unit and during its operation, it experienced two faulty electronics boards, 2) data completeness was only 60% over the one month of operation, and 3) the collected data were extremely noisy.

Table 3-22. Periods of at least 12 consecutive hours of missing sulfate data.

Site	Start Time (PST)	End Time (PST)	Cause
BAC	1/20/2001 0:40	1/22/2001 9:10	Flash Strip Failure
BAC	2/1/2001 20:40	2/2/2001 11:00	Flash Strip Failure

3.6.12 Particle Counts by Climet OPC

Five-minute particle counts from the sixteen size bins (see Table 2-13) of the Climet were reviewed with SurfDat. The parameters reported to ARB include particle concentrations (#/cm^3) from the sixteen channels and sample flow. Prior to validation, all data, missing data excepted, had a QC code of zero (0). As validation progressed, validation actions were limited to data that had QC codes of zero, in order to avoid overwriting other non-zero QC codes. The following steps were taken to validate the Climet OPC data

- Compiled maintenance and flow check summary from off-line summary, log sheets, and task sheets.
- Documented periods of greater than 12 hours of missing data noted in off-line summary and/or logbook. Briefly reviewed data in SurfDat to verify this information and to identify undocumented periods of missing data.

- Documented major repairs or modifications to the instrument that would have affected data (collected before or after the repairs/modification), such as alignment of the laser diode.
- Documented start and end date/time of sampling; documented instrument serial number(s) and if/when one instrument was exchanged for another.
- Completed the set of QC checks listed in **Table 3-23**. Referred to site/instrument logbook, task sheets, and off-line summary to determine appropriate periods over which to invalidate data. Note that all data for the two largest size bins were flagged as "wholly or partially compromised" (QC code=1).
- Neither flow nor concentration correction factors were applied to the data. If flow checks did not pass, affected data and potentially affected data were flagged as suspect or invalid.

Table 3-23. Validation criteria for Climet OPC. Table 2-13 provides bin size cuts.

Parameters to Flag in a Record	Reason for Flag	QC Code	QC Definition	Comments
Flow, all sixteen bins	Documented instrument problems	3	Invalid	Repairs, power failure, etc.
Flow, all sixteen bins	Routine maintenance	2	Invalid	Flow check, dynamic zero, PSL check, etc.
Flow, all sixteen bins	Angiola tower instruments not at correct height	5	Invalid	Applies to only 50-m and 100-m Climet OPC data
Flow, all sixteen bins	Flow rate >±20% 1 LPM ($Q > 1.2$; $Q < 0.8$)	3	Invalid	Flow path obstruction, operator interference, instrument error
Flow, all sixteen bins	Flow rate >±10% 1 LPM ($Q > 1.1$; $Q < 0.9$)	7	Suspect	Flow path obstruction, operator interference, instrument error
Smallest bin (1)	Failed dynamic zero	7	Suspect	Laser diode not aligned correctly; two particles mistaken for one, perhaps
Largest bin (16)	Roof inlet flow >±10% of set point	7	Suspect	Size-cut of particles in air stream adversely affected
Two largest bins (16 and 15)	Inlet is sharp and always removes a fraction of larger particles	1	Data wholly or partially compromised	N. Kreisberg, instrument expert at ADI, recommended this flag

- All data were divided by (flow rate*sample time*1000) to obtain particle concentration in #/cm^3. Sample time was 4 minutes and 45 seconds (4.75 min), and valid flow rates were approximately 1.0 LPM.

- All particle concentrations were rounded to the nearest 5-minute interval; for example, a record that had a modified time stamp of 554 was rounded to 555.
- Data were compared with concurrent nephelometer and Lasair OPC data.
- Data were reviewed by the measurement expert.

Issues with specific instruments are discussed below by site.

Angiola Ground Level

A Climet OPC (S/N 990247) was installed in the trailer at the Angiola site from the end of March 2000 through mid-February 2001. All data from setup on March 30, 2000, until May 9, 2000, were invalidated; during this period, the instrument was returned to the manufacturer for repairs. Although no flow check was documented for this instrument until July 25, 2000, the July 25 flow check was valid, and the instrument displayed consistent, in-range flow from May 9 through July 25, 2000. Some data in mid-May were flagged as suspect because of slightly elevated instrument-indicated flow rates. In general, the only incorrect flow that would affect data was the flow through the roof inlet, which was detected to be low in late July 2000.

The initial Climet OPC installed in the trailer was replaced by another Climet OPC (S/N 990246) in early August 2000. Both ground Climet OPCs passed all polystyrene latex (PSL) checks. However, the second instrument (S/N 990246) failed a few dynamic zero checks.

The Climet OPC S/N 990246 was installed in the 50-m tower briefly (December 6-13, 2000) because the Climet OPC that had been at the 50-m height developed significant problems. This move was critical because no other particle-sizing instrumentation was in operation at that height on the tower. Thus, there were no ground-level Climet OPC data during this period.

Following is a discussion of observations and flagged data:

- In the Climet OPC, channels 3 (0.56-0.69 µm), 4 (0.69-0.86 µm), and 5 (0.86-1.1 µm) were not distinctive. Their counts overlapped and appeared to intermingle considerably. For example, although channel 1 counts were consistently higher than channel 2 counts, channel 2 counts were consistently higher than channel 3 counts, and so on; the channel 5 counts often exceeded those of channels 3 or 4. Similarly, sometimes channel 4 counts exceeded those of channel 3. While this phenomenon is not necessarily expected, it is also not of great concern, as the polystyrene latex checks of bin size consistently passed, and the phenomenon was observed in both ground-level Climet OPC instruments.
- There were substantial baseline shifts downward during the periods October 9 through 13, 2000, October 20 through the end of October 2000, January 6 through 13, 2001, and January 23 through 27, 2001. Though these occurrences all appear to be coincident with changes in meteorology, the period in early January (6-13) occurs immediately after a significant (3.5-hr) break in the data. There is no documentation (e.g., record of maintenance or power outage) explaining the break. These data were not flagged.
- Vacillations (of varying periodicity) in the particle counts occurred November 7 through November 13, 2000. In addition, the baseline was somewhat depressed during this

period. After comparison with Lasair OPC and nephelometer data, the data collected during this period were flagged as invalid.

- A general source of uncertainty in these data, as well as those of the PMS Lasair, is the real, although small (on the order of 1%), probability of two particles being counted as one. If two particles pass through the laser beam simultaneously, they will be counted as one but will scatter more light than either one by itself. Thus this "single" particle will also be classified as larger than either particle actually is. This general phenomenon results in slight undercounting in the small bins and slight overcounting in the large bins.

Angiola 50-m Tower

A Climet OPC (S/N 978182) operated on the 50-m tower at the Angiola site from mid-August 2000 through early February 2001. The primary problems with this instrument were poor flow checks and failed dynamic zeroes

- A high flow rate through the roof inlet results in a decreased cut-point for particles; a low flow rate through the roof inlet results in a larger size cut-point. During a flow check on October 19, 2000, the flow through the inlet was determined to be 48% low. Thus, the operator raised the flow rate. However, an independent flow audit performed a few days later showed that the flow through the inlet was 36% high. Documentation is unclear as to whether the flow was manually adjusted at that time. The next flow check at the inlet, performed on January 18, 2001, showed that the flow through the inlet was flagged as acceptable. All data from the two largest particle size bins collected between October 19 and 23, 2000 were flagged as invalid.
- No flow checks on the instrument were recorded after October 23, 2000.
- No data were recorded December 19, 2000, through midday January 11, 2001. Documentation is unclear as to why data are missing.
- All of the dynamic zeroes from November 7 to December 27, 2000, failed. These failed zeroes were presumed the result of a misaligned laser diode, which was diagnosed and repaired by the site operators. Data in the smallest bin were flagged as suspect (see Tables 3-21).
- There were failed PSL checks in which the highest counts of all size cuts were found in the bin (~0.3-0.4 μm). This failure was also presumed the result of a misaligned laser diode. All data in the smallest bin were flagged as suspect.

Angiola 100-m Tower

A Climet OPC (S/N 990247) was installed on the 100-m tower at the Angiola site from mid-August 2000 through early February 2001. Observations and problems included the following

- Data were recorded at uneven intervals (of one to three hours from startup) through midday on August 31, 2000. All data reported at uneven intervals were invalidated because the sample times and, thus, concentrations, could not be precisely determined.
- All PSL checks on this instrument passed.

- This instrument failed a few dynamic zeroes. Thus, all data in bin 1 were flagged as suspect for the period October 31, 2000 through January 11, 2001.

3.6.13 Particle Counts by PMS Lasair OPC

Five-minute particle counts from the eight size bins (see Table 2-14) of the PMS Lasair OPC were reviewed with SurfDat. The parameters reported to ARB include particle concentrations (#/cm^3) from the eight channels and sample volume. Prior to validation, all data, missing data excepted, had a QC code of zero (0). As validation progressed, validation actions were limited to data that had QC codes of zero in order to avoid overwriting other non-zero QC codes.

The steps taken to validate the PMS Lasair OPC data were the same as those for the Climet OPC as discussed in Section 3.6.12. The QC checks specific to the PMS Lasair OPC are summarized in **Table 3-24**

Table 3-24 Validation criteria for PMS Lasair OPC. Table 2-14 provides bin size cuts.

Parameters to Flag in a Record	Reason for Flag	QC Code	QC Definition	Comments
Volume, all eight bins	Documented instrument problems	3	Invalid	Repairs, power failure, etc.
Volume, all eight bins	Routine maintenance	2	Invalid	Flow check, dynamic zero, PSL check, etc.
Volume, all eight bins	Sample volume >±50% 0.134 L (V > 0.067; V < 0.202	3	Invalid	Flow path obstruction, operator interference, instrument error
Volume, all eight bins	Sample volume >±10% 0.134 L (V > 0.148; V < 0.121)	7	Suspect	Flow path obstruction, operator interference, instrument error
Five smallest bins (1, 2, 3, 4, 5)	Particle count = 0	3	Invalid	Effectively impossible data; operation problem
Bin 6	Particle count = 0	7	Suspect	Unlikely physical phenomenon

The Lasair OPC operated at the Angiola site from late March 2000 until mid-February 2001. From the time that the Lasair OPC was set up in late March 2000 until April 17, 2000, the data were reported to the DAS in one-minute intervals. During this period, the data were flagged as invalid because the operators were experimenting with different sample stream configurations. Following are other observations and problems:

- The Lasair OPC performed well from late April through mid-July 2000. On July 11, 2000, the data began vacillating (with a periodicity of approximately 45 to 60 minutes); this trend continued through July 26, 2000. These data did not track with the contemporaneous nephelometer or Climet OPC data and were flagged as invalid. Toward the end of July 2000, the instrument-indicated sample volume began to drop out of range (±10%). (The sample volume had been slowly declining since mid-June.) While the data track with nephelometer and Climet OPC data, the particle concentrations are not optimum. On July 31, 2000, a flow check confirmed that the sample volume was indeed lower than the set point but not as low as the instrument indicated.

- The flow was adjusted (raised) in early August 2000 but quickly began to drop again. The flow adjustment appeared to have little effect on the particle counts, however. All flows were within 50% of the set point—for this instrument, the ratio of sample flow to sheath flow (not sample flow alone) is the critical parameter. However, instrument-indicated flow was used to calculate particle concentrations, and thus an incorrect indicated flow renders the calculated particle concentrations imprecise. Therefore, all data from August 3 through August 23, 2000 (when the instrument was shipped back to PMS for repair) were flagged as invalid.
- The instrument was returned to the site on September 11, 2000 but was again sent to PMS on September 29, 2000. No flow checks were performed during September 11-29, and the altitude setting of the instrument was not noted. Thus, while these data were not flagged, we recommend they be regarded cautiously.
- The Lasair OPC was returned from PMS on October 19, 2000. From setup on October 19 through midday November 13, 2000, the Lasair OPC sampling interval was approximately 90 minutes. These data were invalidated.
- The Lasair OPC's altitude setting was set to 5850 feet during its time at PMS (the Angiola site has an elevation of approximately 300 feet msl); the flow rate of the instrument on return was about 40% above set point. This discrepancy was not noticed, however, until January 10, 2001. The incorrect elevation setting may have contributed to the flow problems of this instrument in November, December, and early January. During these months, all data were flagged as suspect.
- On January 10, 2001, the improper altitude setting was detected, and on January 11, the correct setting was entered. The flow was also adjusted, and from January 11 through the end of the study, the flow and altitude settings were correct.

3.6.14 Particle Concentrations by SMPS

Five-minute particle concentrations from the 51 size bins (see Table 2-12) of the SMPS were reviewed by ADI. The parameters reported to ARB included particle concentrations (#/cm^3) from the 51 channels and sample flow. The following steps were taken to validate the SMPS data:

- Compiled maintenance and flow check summary from off-line summary, log sheets, and task sheets.
- Documented periods of greater than 12 hours of missing data noted in off-line summary and/or logbook.

- Documented major repairs or modifications to the instrument that would have affected data (collected before or after the repairs/modification).
- Documented start and end date/time of sampling.

The OPC/SMPS site log, task sheets and off-line summary were scanned for information on both regular maintenance, as well as instrument problems including the OPCs, which shared an inlet system with the SMPS. This was done multiple times by both STI and ADI personnel and the data were flagged appropriately. SMPS operating parameters (flows, pressures, etc.) were screened and flagged for out of range values as well as sudden jumps in values from one sample. Generally, useful data were delivered for periods of stable operating conditions (i.e., no maintenance activities, audits, etc.) with flags for conditions such as PSL calibrations, dynamic zeroes, contaminated sample, etc.

Maintenance of the SMPS consisted of a 3-day schedule, a roughly monthly schedule, and periodic actions. The 3-day maintenance list should be doable within one 5-minute sample period, but instrument logs indicated that several samples were usually lost. Routine maintenance resulted in loss of SMPS data as discussed next:

- The TSI SMPS program had to be stopped and restarted as noted in Section 2.5.12 every third day (or sooner).
- The sheath and sample flows were stopped and the sheath flow and impactor pressure drop transducers were re-zeroed to compensate for any zero drift every third day (or sooner).
- The impactor collection plate was cleaned of deposited aerosol and re-greased to prevent particle bounce every third day (or sooner). This minimized accumulated aerosol deposit which gradually increases the pressure drop across the impactor.
- The impactor orifice was also cleaned, but not during every 3-day maintenance check. Orifice cleaning took several sample periods.
- The butanol reservoir in the CPC was refilled (approximately weekly). Usually this was done as part of the 3-day maintenance, however, sometimes the refilling was problematic and took several sample periods.
- Flows checks of the DMA sheath and CPC sample flows were performed as a part of routine maintenance performed monthly. The sheath flow control system is very sensitive to pressure drop and thus, most flow standards (e.g., BIOS) other than a Gilibrator cannot be used. Early in the study a Gilibrator was not available at the site but, even after one was obtained, audits of these flows were only performed twice by site personnel (according to the logs) on July 31, 2000 and August 29, 2000 and twice by ADI personnel during site visits on December 15, 2000 and February 16, 2001. These flow checks, as well as other audits of the SMPS-OPC flow system, took the SMPS system off-line for one to four hours apiece.
- Dynamic zero concentration checks with an inlet filter and particle size calibration checks with monodisperse polystyrene latex (PSL) spheres were performed periodically (roughly

once a month) on the SMPS and OPCs. These checks typically took the SMPS system off-line for about two hours apiece.

Table 3-25 summarizes the data gaps greater than 12 hours in duration.

Table 3-25. Data gaps in the SMPS data greater than 12 hours.

Start Date Time (PST)	End Date Time (PST)	Duration (hours)	Cause
4/21/2000 23:55	4/24/2000 11:34	59.66	Unknown, occurred on a weekend
4/30/2000 20:31	5/1/2000 11:18	14.78	Unknown, occurred on a weekend
5/14/2000 21:21	5/16/2000 9:34	36.22	Reached 999-sample limit, occurred on a weekend
6/8/2000 13:56	6/12/2000 14:24	96.47	Impactor pressure drop too high, impactor dirty?
6/14/2000 14:40	6/15/2000 10:39	20.00	Impactor pressure drop too low, bad zero?
6/20/2000 17:40	6/21/2000 7:34	13.91	Unknown
6/30/2000 10:00	7/10/2000 16:29	246.48	No sheath flow measurements between zeroes, program freeze
7/11/2000 13:39	7/27/2000 12:10	382.51	CPC flooded, replace pressure transducer
8/6/2000 0:31	8/7/2000 10:40	34.14	Reached 999-sample limit, occurred on a weekend
8/9/2000 8:25	8/10/2000 13:25	29.00	CPC flooded
8/26/2000 23:21	8/28/2000 8:45	33.39	Reached 999-sample limit, occurred on a weekend
9/24/2000 23:26	9/26/2000 7:41	32.26	Power outage
10/4/2000 19:15	10/5/2000 9:19	14.07	Unknown
10/12/2000 23:55	10/16/2000 8:24	80.49	Unknown, occurred on a weekend
10/22/2000 18:46	10/23/2000 7:06	12.34	Reached 999-sample limit, occurred on a weekend
10/27/2000 10:16	10/31/2000 9:40	95.40	No sheath flow measurements between zeroes
11/8/2000 16:25	11/9/2000 11:01	18.59	Unknown
12/1/2000 17:30	12/4/2000 8:41	63.19	Unknown, occurred on a weekend
1/1/2001 12:25	1/2/2001 9:04	20.66	Unknown
1/2/2001 16:11	1/3/2001 9:03	16.87	Inlet disconnected
2/11/2001 3:51	2/12/2001 11:50	31.98	Reached 999-sample limit, occurred on a weekend

Specific data processing procedures are described below.

Computing Flow Rate

In order to use the bypass flow transducer data as a measure of the sheath flow, a calibration was needed of actual sheath flow versus bypass flow transducer reading and temperature. The properly zeroed normal sheath flow transducer was used as the flow standard because previous calibrations against a Gilibrator showed it to be accurate. Special software was written to scan the electrostatic classifier (ESC) through a series of sheath flow settings (~5 minutes per setting) while reading the temperature, pressure and flow parameters every 10 seconds. This was run overnight for several days at ADI to obtain the calibration at a range of temperatures. The bypass flow readings were then fitted to a quadratic function in sheath flow with coefficients linear in temperature. Inversion of this equation gave the desired calibration equation. This was used to convert the Angiola ESC bypass flow readings to sheath flow. Because these values were rather noisy, they were averaged to a single value for each period between transducer zeroes. These are the sheath flow values that were used in the final data reduction. They agree to within less than 1% with Gilibrator readings of six sheath flows ranging from 7.0 to 9.6 L/min with independent zeroes measured during four audits over the course of the study. Unfortunately, some periods had no readings of ESC parameters because the SMPSDAT program was stopped. No data has been delivered for these periods because there is no reliable information about the sheath flow rate, which is critical to the sizing of the particles.

Correcting Pressure Errors

The pressure upstream of the impactor is determined as the DMA absolute pressure plus the pressure drop across the impactor. The necessary continuity of the upstream pressure through transducer zeroing and impactor cleaning was used to determine the offset of the impactor pressure drop transducer when it was improperly zeroed and make the appropriate corrections.

Accounting for SMPS Program Problems

Linear interpolations were used to fill in the gaps in the required operating parameters (besides sheath flow) resulting from the freezing of the SMPSDAT program. To a close approximation, the impactor pressure drop increased linearly over time between impactor cleanings while the DMA absolute pressure decreased correspondingly. The DMA temperature and the impactor upstream pressure both showed some degree of diurnal variation. However, the neglect of this on interpolations covering a substantial fraction of a day to several days was determined to have insignificant effects on calculations of the final particle size distributions.

Correcting File Dates and Operating Parameters

The first step in post-processing the Angiola SMPS data was to correct the erroneous dates on the raw sample files. This was largely accomplished by comparing the sample dates in the serial data sent to the DAS to the date of the daily serial data capture file in which they were stored. All SMPS data were then reprocessed to correct for the problems with transducer zeroing and the gaps in the operating parameter data.

Converting Counts to Concentration

The SMPS raw data consist of particle count rate versus scan time. This must be converted to particle concentration versus particle diameter. Theoretical relationships are used to convert from scan time to particle electric mobility to particle diameter. These calculations take into account DMA dimensions, voltage scan rate and limits, sheath and aerosol flow rates and flow transit time from the DMA to the CPC. They also depend on the temperature and absolute pressure in the DMA. These last two are measured operating parameters and the sheath flow is determined from the bypass flow reading and the temperature. The remaining parameters were fixed throughout the study.

Particle count rate is converted to particle concentration by dividing by the sample flow rate and the charging, DMA transmission, and CPC detection efficiencies. Size-dependent aerosol charge fractions are determined theoretically and are provided by TSI. DMA transmission efficiency is determined by the aerosol to sheath flow ratio. Size-dependent CPC detection efficiencies are provided by TSI and have been determined experimentally for this model CPC but not for this particular unit.

The SMPS sample flow is determined by the 1 L/min critical orifice at the exit of the CPC. The actual sample flow is calculated as the volumetric flow upstream of the impactor. The main pressure drop between this point and the orifice is the pressure drop across the impactor. The sample flow is therefore calculated as the flow rate downstream of the impactor (1 L/min) multiplied by the ratio of downstream (DMA) to upstream absolute pressures.

Correcting for Multiply-Charged Particles

Correction of the final size distribution for multiply-charged particles exiting the charger is also possible. This essentially amounts to subtracting off the contribution of multiply-charged particles from the size distribution. However, this correction is of limited accuracy largely due to variation of charger performance from theory but also due to variations of the sampled size distribution during the course of the 5-minute combined scans and Poisson-limited count rate measurement accuracy. These inaccuracies can lead to meaningless negative particle concentrations.

The effect of the correction is greatly dependent on the shape of the size distribution. For the Angiola SMPS data set, the effect of the correction was too significant to ignore. The multiple-charge correction was applied to all the data and any negative concentrations so generated were set to zero concentration. The correction only affected size channels 11 to 45 or 20.5 to 255 nm particle diameter. The impactor cuts off larger sizes that would affect channels above 45 and multiple-charge fractions become insignificant below channel 11. Within these 35 size channels containing 72,136 valid or suspect ambient distributions, this correction generated 1678 false zeroes.

The data for the two smallest size channels (8.66-10 nm) obtained from the SMPS are compromised to such an extent that they have not been reported in the final data. Below 10 nm the correction factor for particle concentration due to the drop off in CPC detection efficiency is greater than 3. But there is also a significant uncertainty in this factor because the detection efficiency for this particular unit was not measured.

Correcting for "Time-Smearing"

The lowest channels of the size distribution can be significantly affected by "time-smearing" of the aerosol. Particles leaving the DMA at the same time have a distribution of transit times to the CPC due to the parabolic velocity profile of the laminar flow through the transit tube. This effectively smears the measured distribution of count rate versus scan time over the time parameter. This affects the entire scan but the most prominent effects occur where relatively low count rates are immediately preceded by much higher count rates. Time smearing then artificially raises the count rate in the low count zones. The lowest size channels are measured at the beginning of the DMA voltage up scan and the count rates are generally low because of the drop off of CPC detection efficiency. But time smeared remnants from the preceding rapid down scan may be of notably larger particle sizes with much higher detection efficiencies in the CPC. This can result in a significant positive artifact in these lowest channels.

It is difficult to correct for the effects of time smearing without greatly magnifying the already present noise in the Poisson counting statistics. The SMPS data have not been corrected for this. Instead, the bottom two channels of the size distribution have not been reported since they suffered the most from time smearing. The lowest reported channels are still somewhat affected but to a much lesser degree. Care should be taken in inferring too much from a small upturn in reported concentration at the lower end of the distribution if it only involves the lowest two or three reported channels.

Validation Results at Angiola

There were several problems at the Angiola site that seriously impacted the completeness of the SMPS data. The most important of these were improper flow transducer zeroing and freezing of the SMPSDAT program. Occasional faulty computer date recovery at power up, out of range impactor pressure drop, CPC butanol flooding, and exceedance of the 999-sample limit caused data loss as well.

Before zeroing the flow and pressure drop transducers, the sheath and sample flows must be stopped. The latter is stopped by the operator turning off the vacuum pump downstream of the CPC while the sheath flow is automatically set to zero upon entering the transducer-zeroing screen on the front panel of the ESC. The sample pump stops immediately and the relatively small flow volume between the CPC critical orifice and the pump reaches equilibrium pressure fairly quickly to stop the actual sample flow. The sheath flow, on the other hand, is driven by a pair of in-line blowers, which actually take 10-20 seconds to spin down after the power is removed. If the sheath flow transducer is zeroed while there is still residual flow, then, in normal operation, the transducer reading will be low when the correct sheath flow (7 L/min) is flowing. Since the ESC uses this same transducer in the control loop for the blowers, it will now drive the blowers to higher flow until the correct transducer reading (7 L/min) is obtained.

The sheath flow transducer was frequently improperly zeroed in this manner at the Angiola site resulting in higher than expected sheath flows and under-sizing of the aerosol selected by the DMA. Fortunately, there was a second flow transducer in the sheath flow line for the ESC configuration used at Angiola. Since there was no "bypass" flow needed for the long DMA, the bypass flow blower and transducer were configured in series with the normal sheath flow blower and transducer. The two blowers are identical but the bypass flow transducer is simply an orifice with a pressure drop transducer compared to the more accurate laminar flow transducer used to control the sheath flow. The natural order of flow transducer zeroing for the ESC is first sheath flow, then bypass flow, and then impactor. So, though there may have been residual sheath flow when the sheath flow transducer was zeroed, the flow was generally stopped by the time the bypass flow transducer was zeroed. Improper zeroes happened frequently enough that it has been assumed that the sheath flow transducer readings generally cannot be trusted but that the bypass flow transducer readings represent a consistent measure of the sheath flow.

For unknown reasons the SMPSDAT program temporarily stalled or completely froze 61 times during the study at Angiola resulting in delayed processing of 11.4% of the SMPS samples in all. Temporary stalls generally only affected a few samples at each occurrence but program freezes lasted anywhere from several hours to several days before an operator would restart the program. In one case, the SMPSDAT program was stopped for five days allowing more than 1400 sample files to pile up in the SMPS data directory. This resulted in freezing of the SMPS program for another five days. The SMPSDAT program was written such that it realized when it was delayed picking up the SMPS data file, and in such case, it did not read the ESC parameters because those readings would not be at the same time the count data were taken. Instead, the program used the last set of parameter readings to reduce the data.

There were ten periods, affecting 5.2% of the samples, during which the SMPS computer date was wrong while the time was correct. For most of these periods, the erroneous date (usually January 1980) was incorrectly recovered from the battery backup at computer power up after a power outage or after being shut down by the operator. However, one period began in the middle of a sample when the computer date spontaneously changed from the correct date to January 4, 1980. The SMPS raw sample files are renamed by the SMPSDAT program according to the sample start date and time – Amddhhmm.DAT. As a result of the date recovery problem, files names were erroneously duplicated 43 times causing data overwrite resulting in the loss of 43 sample files when the files were zipped and deleted.

There were two periods, affecting 2.3% of the samples, during which the pressure drop across the impactor read unusually low while the DMA absolute pressure read normally. The abnormal readings were probably a result of improper zeroing of the pressure drop transducer. Since this reading is not used as a control parameter for the SMPS, this problem did not affect the operation of the instrument.

Another 3.2% of the samples had out of range impactor pressure drop (with corresponding deviations in DMA absolute pressure) for a variety of reasons. The main reason was impactor overload. It is possible that unrecorded operator activity accounts for many brief periods of deviation. About 1.0% of these cases were extended periods of marginally out of range values; data were flagged as suspect. The remaining 2.2% of the cases were invalidated.

The CPC optics and other parts were flooded with butanol twice, apparently due to problems with the shutoff valve during filling. In one case the vacuum pressure sensor was damaged requiring a replacement. These problems accounted for 411.5 hours (~17 days) of downtime. The 999-sample limit was exceeded 16 times accounting for 241.5 hours (~10 days) of downtime.

Tables 3-26 and 3-27 list the primary and secondary flags used during data validation and reported to ARB. Note that when a sample was flagged, the flag was applied to all data channels of that sample.

Table 3-26. Primary flags used by ADI during data validation of the SMPS data.

Primary Flag[a]	Description
I	invalid ambient data
S	suspect ambient data
V3 > M	valid ambient data, raw data lost
V2	valid ambient data, no original operating parameters, operating parameters interpolated where possible
V6 > V2	valid ambient data, non-standard sampling setup
V0	valid ambient data

[a] In order of decreasing priority.

Table 3-27. Secondary flags used by ADI during data validation of the SMPS data.

Secondary Flag[a]	Primary Flag	Description
ZT7	I[b]	zero flow transducers: sheath, bypass, impactor
ZT3	I[b]	zero flow transducers: sheath, bypass
ZT6	I[b]	zero flow transducers: bypass, impactor
ZT1	I[b]	zero flow transducers: sheath
ZT2	I[b]	zero flow transducers: bypass
ZT4	I[b]	zero flow transducers: impactor
M3D	I[b]	3-day maintenance
CSH	I	change sheath flow setting
AUD	I	flow audit
CPC	I	CPC off-line
PSL	I	PSL calibration
ZER	I	zero air
APO	I	absolute pressure out of range
IPO	I,S	impactor pressure drop out of range
ISH	I,S	improper sheath flow setting
CTM	I,S	contamination other than PSL or room air
RMA	I,S	room air contamination
USH	S	unstable sheath flow
NTF	S	no transport flow
ZO4	S	flow transducer zero offset: impactor
RDL	V3	raw data lost
DPR	V2	delayed processing, no original operating parameters
CL3	V6	three Climets in sampling setup
CL0	V6	no Climet in sampling setup
DAY	V0	incorrect original date
RST	V0	restart sampling

[a] In order of decreasing priority
[b] Except when zeroing or maintenance occurred prior to sampling restart

This page is intentionally blank.

4. GUIDE TO THE ANCHOR SITE DATA ARCHIVE

Data were collected (Wittig et al., 2003), processed, and stored at STI (Section 2), validated (Section 3), and delivered to ARB. In this section, the data delivery process is discussed including a summary of deliverables, the data averaging protocol, and the data reporting format.

4.1 OVERVIEW

Figure 4-1 is a schematic of the data delivery process. After the data were validated (as described in Section 3), validated data were brought back into the SQL database using a procedure built into the SurfDat program. Only the DataValue and QCCode fields of the data record were updated in the SQL database during this procedure (i.e., the RawDataValue was not altered). Comments made and actions taken during validation were also brought into the SQL database.

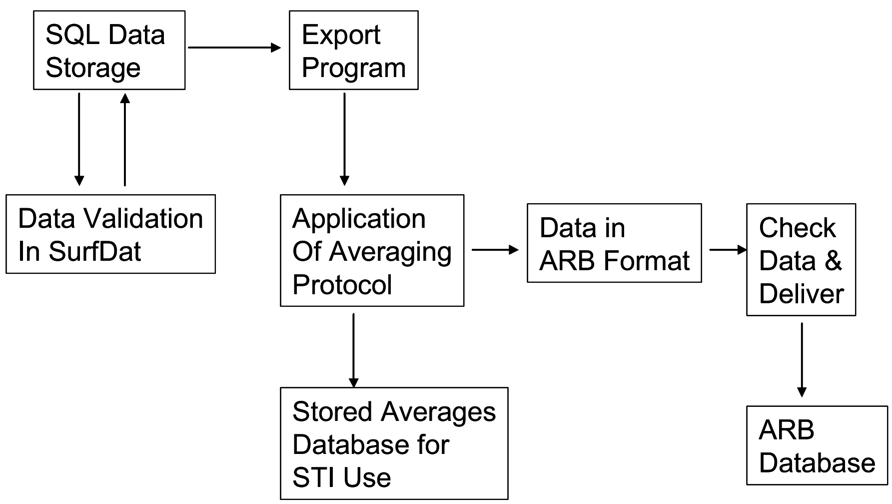


Figure 4-1. Overview of the data delivery process.

Once approval had been given for the delivery of a validated parameter by the measurement expert, the STI technician used SurfDat's export feature to select the appropriate validated records (based on site, parameter, and time period requested) and create a data file with the appropriate format for delivery to the ARB database. The data were averaged according to a protocol developed to address the 1-minute, 5-minute, and 10-minute sampling interval data. The averaged data were then formatted according to ARB instructions. The STI technician followed a comprehensive check list to ensure that the data file was correct.

The files created through the export program were then delivered to ARB by placing the files on an FTP site set up for the exchange by ARB.

4.2 DATA DELIVERABLES

Table 4-1 lists the parameters that were delivered to the ARB. The following data were exported

- Data with a sampling frequency of 1-minute were exported as 5-minute and 60-minute averages[5]
- Data with a sampling frequency of 5–minutes were exported as 5-minute and 60-minute averages[6]
- Data with a sampling frequency of 10-minutes were exported as 10-minute and 60-minute averages
- Data with a sampling frequency of 60-minutes were exported as 60-minute data.

Parameter sampling frequencies were discussed in Section 2. Averaging protocols are described in the next section.

4.3 DATA AVERAGING PROTOCOL

The data deliverables to ARB were contracted to be 5-minute, 10-minute (for minimum sampling intervals of 10 minutes), and 60-minute averages. Existing averaging protocols developed by the ARB and EPA did not specifically address data collected with a sampling frequency less than 60 minutes; these protocols were used to develop averaging protocols for the shorter sample durations.

4.3.1 Overview

Table 4-2 contains information about the number of records required to calculate an average.

Instrument full-scale information (same for all sites) was used to determine whether a 60-minute average could be computed from 30 to 44 minutes of data as shown in **Table 4-3**.

Only data having QC codes of 0 (valid), 1 (estimated), 4 (RH>75%, nephelometer-data specific), 7 (suspect), 10 ("averaged 1-minute data", nephelemoter only, valid), or 11 ("slightly suspect—heater on continuously", nephelemoter only, suspect) were used in computing an average. Data accompanied by QC codes 2 (calibration, instrument check), 3 (instrument problem), 5 (Aethalometer tape transfer or tower instrument operating at different height), 6 (failed QC), 8 (invalid), or 9 (missing) had an assigned data value of –980 and were not used in creating averages. However, the reasoning and severity of any data records excluded from an average were retained in order to evaluate secondary flags associated with the final averaged record. **Table 4-4** describes several secondary flags created by STI that illustrate the characteristics of data that comprise a given average. The hierarchy (increasing severity) of assigned QC codes is 0<1<4<7<9<(2, 5, 6)<8; QC code 8 (invalid) is the most severe.

[5] Except PAN data, for which 15-minute and 60-minute averages were reported to ARB.
[6] Except SMPS data, for which only 5-minute data were reported to ARB.

Table 4-1. List of parameters delivered by STI to ARB.

Page 1 of 5

Site	Instrument	Description of delivered parameters	Responsible for QC	Data Start (PST)	Data End (PST)
Altamont	Beta Attenuation Mass Monitor	$PM_{2.5}$ Mass Concentration and volume	STI	1/28/2000 18:00	2/8/2001 8:00
Angiola	Aethalometer (1-wavelength)	Black carbon concentration; instrument flow	STI	1/14/2000 12:35	10/26/2000 8:35
Angiola	Aethalometer (7-wavelength)	Black carbon concentration at 450, 571, 590, 660, 880, and 950 nm; instrument flow	STI	10/26/2000 12:50	2/28/2001 23:55
Angiola	Ambient Carbon Particulate Monitor (OCEC)	Organic carbon and elemental carbon mass concentration and volume	STI	2/23/2000 7:00	2/9/2001 23:00
Angiola	Ambient Particulate Nitrate Monitor	Nitrate concentration, ambient NO_x concentration	ADI	11/21/2000 0:00	3/2/2001 9:20
Angiola	Beta Attenuation Mass Monitor	PM_{10} Mass Concentration and volume	STI	1/20/2000 8:00	2/6/2001 10:00
Angiola	Beta Attenuation Mass Monitor	$PM_{2.5}$ Mass Concentration and volume	STI	1/20/2000 7:00	2/6/2001 8:00
Angiola	Climet Optical Particle Counter	Particle counts for 16 size-cut bins; sample flow	STI	3/30/2000 14:40	2/8/2001 20:55
Angiola	Integrating Nephelometer	Light scattering by particles (b_{sp}); relative humidity, pressure, and temperature of instrument	STI	2/1/2000 14:00	2/9/2001 23:55
Angiola	Nitric Acid Analyzer	NO_y and NO_y-HNO_3 concentrations	STI	12/13/2000	2/4/2001 23:59
Angiola	NO_2/PAN Gas Chromatograph	NO_2 concentration	CE-CERT, STI	11/19/2000 14:49	2/12/2001 13:00
Angiola	NO_2/PAN Gas Chromatograph	PAN concentration	CE-CERT, STI	12/26/00 0:15	1/7/01 23:45
Angiola	NO_y Analyzer	NO concentration, NO_y concentration	STI	2/9/2000 0:00	2/5/2001 6:00
Angiola	Ozone Analyzer	Ozone concentration	STI	1/22/2000 0:00	2/21/2001 2:45
Angiola	PMS Lasair Optical Particle Counter	Particle counts for 8 size-cut bins; sample volume	STI	3/30/2000 14:40	2/16/2001 20:55
Angiola	Scanning Mobility Particle Sizer (SMPS)	Particle counts for 53 size-cut bins; sample flow	ADI	4/3/2000 12:15	2/16/2001 19:55
Angiola Tower 1 m	Integrating Nephelometer	Light scattering by particles (b_{sp}); relative humidity, pressure, and temperature of instrument	STI	12/14/2000 19:45	2/9/2001 23:55

Table 4-1. List of parameters delivered by STI to ARB.

Site	Instrument	Description of delivered parameters	Responsible for QC	Data Start (PST)	Data End (PST)
Angiola Tower 50 m	Climet Optical Particle Counter	Particle counts for 16 size-cut bins; sample flow	STI	8/18/2000 13:00	2/12/2001 15:00
Angiola Tower 50 m	Integrating Nephelometer	Light scattering by particles (b_{sp}); relative humidity, pressure, and temperature of instrument	STI	8/18/2000 13:20	2/9/2001 23:55
Angiola Tower 100 m	Aethalometer (7-wavelength)	Black carbon concentration at 450, 571, 590, 660, 880, and 950 nm; instrument flow	STI	12/1/2000 16:20	2/16/2001 17:00
Angiola Tower 100 m	Ambient Particulate Nitrate Monitor	Nitrate concentration, ambient NO_x concentration	ADI	12/5/2000 19:30	3/2/2001 9:20
Angiola Tower 100 m	Climet Optical Particle Counter	Particle counts for 16 size-cut bins; sample flow	STI	8/18/2000 15:40	2/16/2001 12:10
Angiola Tower 100 m	Integrating Nephelometer	Light scattering by particles (b_{sp}); relative humidity, pressure, and temperature of instrument	STI	8/18/2000 0:00	2/9/2001 23:55
Angiola Tower 100 m	NO_y Analyzer	NO concentration, NO_y concentration	STI	12/5/2000 0:00	2/5/2001 0:05
Angiola Tower 100 m	Ozone Analyzer	Ozone concentration	STI	12/4/2000 22:10	2/8/2001 23:55
Bakersfield	Aethalometer (1-wavelength)	Black carbon concentration; instrument flow	STI	1/20/2000 0:00	10/23/2000 13:25
Bakersfield	Aethalometer (7-wavelength)	Black carbon concentration at 450, 571, 590, 660, 880, and 950 nm; instrument flow	STI	10/23/2000 16:30	2/14/2001 23:55
Bakersfield	Ambient Carbon Particulate Monitor (OCEC)	Organic carbon and elemental carbon mass concentration and volume	STI	10/4/2000 11:00	2/9/2001 23:00
Bakersfield	Ambient Particulate Nitrate Monitor	Nitrate concentration, ambient NO_x concentration	ADI	11/15/2000 12:30	3/6/2001 10:00
Bakersfield	Ambient Particulate Sulfate Monitor	Sulfate concentration, ambient SO_2 concentration	ADI	1/13/2001 12:50	2/16/2001 8:50
Bakersfield	Beta Attenuation Mass Monitor	PM_{10} Mass Concentration and volume	STI	1/21/2000 23:00	2/6/2001 15:00
Bakersfield	Beta Attenuation Mass Monitor	$PM_{2.5}$ Mass Concentration and volume	STI	1/21/2000 23:00	2/6/2001 15:00

Table 4-1. List of parameters delivered by STI to ARB.

Site	Instrument	Description of delivered parameters	Responsible for QC	Data Start (PST)	Data End (PST)
Bakersfield	Integrating Nephelometer	Light scattering by particles (b_{sp}); relative humidity, pressure, and temperature of instrument	STI	1/6/2000 17:10	2/9/2001 23:55
Bakersfield	NO_2/PAN Gas Chromatograph	NO_2 concentration	CE-CERT, STI	10/11/2000 10:20	2/12/2001 9:34
Bakersfield	NO_2/PAN Gas Chromatograph	PAN concentration	CE-CERT, STI	12/26/00 0:15	1/7/01 23:45
Bakersfield	NO_y Analyzer	NO concentration, NO_y concentration	STI	2/25/2000 17:00	2/16/2001
Bakersfield	SO_2 Analyzer	SO_2 concentration	STI	11/20/2000 19:30	2/9/2001 0:00
Bethel Island	Aethalometer (7-wavelength)	Black carbon concentration at 450, 571, 590, 660, 880, and 950 nm; instrument flow	STI	11/17/2000 16:05	2/15/2001 5:55
Bethel Island	Ambient Particulate Nitrate Monitor	Nitrate concentration, ambient NO_x concentration	ADI	11/28/2000 13:00	2/6/2001 11:30
Bethel Island	Beta Attenuation Mass Monitor	$PM_{2.5}$ Mass Concentration and volume	STI	11/17/2000 13:00	2/15/2001 4:00
Bethel Island	NO_2/PAN Gas Chromatograph	NO_2 concentration	CE-CERT, STI	11/22/2000 14:20	2/12/2001 18:31
Bethel Island	NO_2/PAN Gas Chromatograph	PAN concentration	CE-CERT, STI	12/26/00 0:15	1/7/01 23:45
Bethel Island	NO_y Analyzer	NO concentration, NO_y concentration	STI	11/18/2000 15:00	2/4/2001 16:55
Bodega	Aethalometer (1-wavelength)	Black carbon concentration; instrument flow	STI	11/21/2000 12:50	2/10/2001 18:45
Corcoran	Aethalometer (7-wavelength)	Black carbon concentration at 450, 571, 590, 660, 880, and 950 nm; instrument flow	STI	9/13/2000 11:50	11/14/2000 23:55
Corcoran	Beta Attenuation Mass Monitor	PM_{10} Mass Concentration and volume	STI	9/13/2000 15:00	11/14/2000 22:00
Corcoran	Beta Attenuation Mass Monitor	$PM_{2.5}$ Mass Concentration and volume	STI	9/13/2000 14:00	11/14/2000 22:00
Modesto	Aethalometer (1-wavelength)	Black carbon concentration; instrument flow	STI	10/10/2000 13:40	2/6/2001 12:05
Edwards Air Force Base	Aethalometer (7-wavelength)	Black carbon concentration at 450, 571, 590, 660, 880, and 950 nm; instrument flow	STI	6/20/2000 16:50	9/1/2000 8:45

Table 4-1. List of parameters delivered by STI to ARB.

Site	Instrument	Description of delivered parameters	Responsible for QC	Data Start (PST)	Data End (PST)
Edwards Air Force Base	Beta Attenuation Mass Monitor	PM_{10} Mass Concentration and volume	STI	6/20/2000 19:00	9/1/2000 6:00
Edwards Air Force Base	Beta Attenuation Mass Monitor	$PM_{2.5}$ Mass Concentration and volume	STI	6/20/2000 19:00	9/1/2000 6:00
Sacramento	Aethalometer (1-wavelength)	Black carbon concentration; instrument flow	STI	1/20/2000 0:00	10/6/2000 8:09
Sacramento	Aethalometer (7-wavelength)	Black carbon concentration at 450, 571, 590, 660, 880, and 950 nm; instrument flow	STI	10/6/2000 11:00	2/9/2001 23:55
Sacramento	Beta Attenuation Mass Monitor	$PM_{2.5}$ Mass Concentration and volume	STI	4/13/2000 15:00	2/7/2001 7:00
Sacramento	Integrating Nephelometer	Light scattering by particles (b_{sp}); relative humidity, pressure, and temperature of instrument	STI	12/24/1999 23:00	2/9/2001 23:55
San Jose	Aethalometer (1-wavelength)	Black carbon concentration; instrument flow	STI	1/20/2000 0:00	10/4/2000 10:31
San Jose	Aethalometer (7-wavelength)	Black carbon concentration at 450, 571, 590, 660, 880, and 950 nm; instrument flow	STI	10/4/2000 12:50	2/9/2001 12:30
San Jose	Ambient Particulate Nitrate Monitor	Nitrate concentration, ambient NO_x concentration	ADI	11/9/2000 10:10	2/15/2001 23:40
San Jose	Beta Attenuation Mass Monitor	$PM_{2.5}$ Mass Concentration and volume	STI	5/18/2000 18:00	2/15/2001 21:00
San Jose	Integrating Nephelometer	Light scattering by particles (b_{sp}); relative humidity, pressure, and temperature of instrument	STI	2/3/2000 18:45	2/9/2001 23:55
Sierra Nevada Foothills	Aethalometer (7-wavelength)	Black carbon concentration at 450, 571, 590, 660, 880, and 950 nm; instrument flow	STI	11/19/2000 18:00	2/14/2001 14:20
Sierra Nevada Foothills	Beta Attenuation Mass Monitor	$PM_{2.5}$ Mass Concentration and volume	STI	11/19/2000 13:00	2/12/2001 12:00
Sierra Nevada Foothills	Nitric Acid Analyzer	NO_y and NO_y-HNO_3 concentrations	STI	12/1/2000 18:40	2/15/2001 13:00
Sierra Nevada Foothills	NO_2/PAN Gas Chromatograph	NO_2 concentration	CE-CERT, STI	11/10/2000 13:30	2/13/2001 14:38

Table 4-1. List of parameters delivered by STI to ARB.

Site	Instrument	Description of delivered parameters	Responsible for QC	Data Start (PST)	Data End (PST)
Sierra Nevada Foothills	NO_2/PAN Gas Chromatograph	PAN concentration	CE-CERT, STI	12/26/00 0:15	1/7/01 23:45
Sierra Nevada Foothills	NO_y Analyzer	NO concentration, NO_y concentration	STI	11/24/2000 0:00	2/6/2001 10:20
Sierra Nevada Foothills	Ozone Analyzer	Ozone concentration	STI	12/1/2000 0:00	2/13/2001 14:45
Sierra Nevada Foothills	Ambient Particulate Nitrate Monitor	Nitrate concentration, ambient NO_x concentration	ADI	11/19/2000 20:00	2/6/2001 4:10
Walnut Grove	Aethalometer (1-wavelength)	Black carbon concentration; instrument flow	STI	11/13/2000 12:55	2/13/2001 23:51
Walnut Grove	Ambient Particulate Nitrate Monitor	Nitrate concentration, ambient NO_x concentration	ADI	11/15/2000 17:20	2/13/2001 13:40
Walnut Grove Tower	Aethalometer (1-wavelength)	Black carbon concentration; instrument flow	STI	11/14/2000 17:40	2/13/2001 23:55
Walnut Grove Tower	Integrating Nephelometer	Light scattering by particles (b_{sp}); relative humidity, pressure, and temperature of instrument	STI	11/26/2000 16:15	2/9/2001 23:55
Walnut Grove Tower	Ambient Particulate Nitrate Monitor	Nitrate concentration, ambient NO_x concentration	ADI	11/17/2000 14:00	2/4/2001 0:40

Table 4-2. Sample averaging for STI data delivery to ARB.

Average to be Delivered	Sampling Interval	Percentage	Comments
1 hour	1-minute, 5-minute, or 10-minute	< 50	No average is computed; QC flag indicates insufficient data available
1 hour	1-minute, 5-minute, or 10-minute	≥ 50 and < 75	If the pollutant concentration <25% of the instrument full scale and the maximum value/minimum value <2, the secondary QC flag notes 30 to 44 minutes. Otherwise, no average is computed and the primary QC flag reflects failure.
1 hour	1-minute, 5-minute, or 10-minute	≥ 75 and < 100	Secondary QC flag notes 45 to 59 min.
5 minute	1-minute	< 75	No average is computed; the QC flag indicates insufficient data available
5 minute	1-minute	≥ 75	Secondary QC flag notes whether 4 or 5 valid data points were averaged

Table 4-3. Instrument full-scale information.

Instrument	"Full Scale"
Ozone	500 ppb
NO/NO_y	500 ppb
Nephelometer b_{sp}	10^6 m^{-1}

Table 4-4. STI secondary flags.

Avg Period	Primary Flag	Secondary Flag	Secondary Flag Description
60	V0	A3I	30-44 minute averaged - remainder contains invalid data
60	V0	A3M	30-44 minute averaged - remainder missing
60	V0	A4I	45-59 minute averaged - remainder contains invalid data
60	V0	A4M	45-59 minute averaged - remainder missing
60	I	I30	<30 minute averaged - remainder contains invalid data
60	I	M30	< 30 minute averaged - remainder missing
5	I	MM1	< 4 minute average - remainder missing
5	I	IM1	< 4 minute average - remainder contains invalid data
5	V0	A1I	4 minute average - 1 minute invalid
5	V0	A1M	4 minute average - 1 minute missing

The following is the established protocol for determining the primary (validity) and secondary flag of an averaged data record.

- For the 5-minute average, if any records comprising the average had a QC code of 7 (suspect), the average was assigned a suspect primary flag and a secondary flag that reflected the type of data making up the rest of the average data set (according to Table 4-2).
- For the 60-minute average, if any of the minutes (even 1 minute) within the average had a QC code of suspect, the entire 60-minute average QC code was flagged as suspect. Secondary flags were assigned according to the criteria outlined in Table 4-4.

4.3.2 One-Minute Data

One-minute data were not delivered at the 1-minute time resolution. Rather, 5-minute and 60-minute averages were created. Because data were theoretically available every minute of the hour, the averages were generated by grouping data back to the earliest 5-minute mark (for 5-minute averages) or the 60-minute mark (for hour averages) on the clock. For example, 5-minute averaged data records created from 1-minute records between 0425 and 0429 had a time stamp of 0425 (as start time). Similarly, 60 minutes of data were averaged on the hour, every hour (e.g., 60 1-minute records of data between 1800 and 1859 were averaged together to one record with a time stamp of 1800).

4.3.3 Five-Minute Data

Five-minute data were delivered as the original 5-minute sampling interval and as a 60-minute average. The 60-minute averages were created from all available 5-minute data records occurring within an hour period rather than a fixed number (e.g., 20) of records per average. Therefore, if 5-minute data occurred at 0504, 0509, 0514,…0554, 0559, all data within the 0500-hour period were averaged as an hour record occurring (with a start time) at 0500.

4.3.4 Ten-Minute Data

The 10-minute data records were delivered as the original 10-minute sampling interval and as 60-minute averages. The 60-minute average was created following the same procedure as for that for 5-minute sampling interval data; all 10-minute records available within a 1-hr period (starting at the top of the hour) were averaged together.

4.4 DATA REPORTING (ARB FORMAT)

A significant effort was expended to process data stored in the STI database into ARB's desired format for delivery. The ARB format was documented on the website (http//www.arb.ca.gov/airways/CRPAQS/lookups.htm) and updated occasionally. STI's parameter names, site codes, et cetera, needed to match those in use by ARB. ARB provided data contributors with reference tables. The most frequently used tables included Supports (monitoring stations identified with an ID and code accompanied by name, address, coordinates, elevation, etc.); Parameters (parameter code, descriptions, units, averaging time, applicable temperature and pressure adjustments); Instruments (range, accuracy, and Instrument_ID of instruments used to acquire measurements whether from a station, airplane, blimp or tower); and Methods (essentially a description of how an instrument was utilized to obtain an observation data value). STI worked with ARB to prepare the parameter table.

Three features of the database are noteworthy:

- The values for missing data were reported as a "Null" value accompanied by a validity flag of "MIS".
- Invalidated records were also replaced with null codes and flagged "INV".
- Times were expressed in Pacific Standard Time (PST). Years were reported as four-digit codes (e.g., 1999, 2000, 2001).

STI used an export routine to extract the data and notes from the SQL database, prepare the averages according to the averaging protocol, match parameter names and QC flags to ARB conventions, and format the data according to ARB's requirements. A copy of the data averaged by the export program, specifically designed for export to ARB, was retained in a table format in another database (savedaverages.mdb) automatically. This step was performed because the ARB ASCII format was somewhat unwieldy for STI's future use in-house. Thus, STI has archived the SQL databases for unaveraged data (at the original time resolution) as well as databases containing data averaged by the export program prior to creation of the ARB file format.

Sometimes, there were greater than three notes tied to a single data record. This was prevalent especially when data of small time resolution (e.g., 1-minute or 5-minute data) were averaged, and notes corresponding to individual records were grouped together following the average. Because the ARB format accepted only three notes for each record, a third and any outstanding notes associated with a record were concatenated. For this reason, many of the notes found in the notes section of the ARB file may appear odd; for example, unrelated events may be described in the same note record. This was the only way in which the data could be prepared to meet the restrictions on the quantity of notes allowed by the format. Once the files were created, the STI technician followed a comprehensive check list to ensure that the data file was correct. An excerpt from a check list is provided in **Figure 4-2**.

The files created through the export program were then delivered to ARB by placing the files on an FTP site set up for the exchange by ARB. The export program electronically recorded the data contained within each file to a data table for record-keeping purposes; however, a list of the files that actually were accepted by ARB was manually recorded. An example of an ARB format file is provided in **Figure 4-3**.

After posting the file on the ARB FTP site, we checked the ARB database to ensure that the data had been accepted and that the data posted in the database matched the data files that had been sent. Any problems experienced during the delivery process were resolved with the ARB staff.

Please list: Today's date: __________
ARB File Name: __________ Site: ________ Parameter: ____________ Time Period
From: ________________ to _______________

- ◻ Check that the ARB site id is correct
- ◻ Check that the ARB site text id is correct
- ◻ Check that the ARB Parameter id is correct
- ◻ Check that the ARB method id is correct
- ◻ Check that the ARB method code is correct
- ◻ Check Instrument Tracking List for change of instrument, verify correct number
- ◻ Check that there are "–999" values for Invalid "I" data
- ◻ Check that Missing "M" have blank/null values
- ◻ Check that there are notes corresponding to data values
- ◻ Check that the footer and header is shown
- ◻ Check that the correct Obs_Type_Code is associated with the data
- ◻ Check that the Averaging Interval in the header correctly identifies the data

Figure 4-2. Excerpt of 1-hr average file check list (applied prior to delivery of the data to ARB).

```
1,"ST","F","SFPMMASS","R","20021014","B","S","1A",1735,
5,1,1,
6,1,1,"consecutive identical values 3",
7,1,1,
5,2,1,
6,2,1,"consecutive identical values 4",
7,2,1,
5,3,1,
6,3,1,"instrument check & maintenance; flow audit",
7,3,1,
5,4,1,
6,4,1,"negative concentration value",
7,4,1,
5,5,1,
6,5,1,"PM10=0 ug/m3; Volume=0 L",
7,5,1,
5,6,1,
6,6,1,"PM2.5>PM10",
7,6,1,
8,"",207,"EDW","20000715","20000715",0,0,0,"PST","03:00:00","04:00:00",416,417,"MASS
_PAR_PU0000010000_BAM_BTD_HD_H1",124,"V0"," ","",33,,"HD",,,,,,,"","",
8,"",207,"EDW","20000715","20000715",0,0,0,"PST","15:00:00","16:00:00",416,417,"MASS
_PAR_PU0000010000_BAM_BTD_HD_H1",124,"V0"," ","",32,,"HD",,,,,,,"","",
8,"",207,"EDW","20000714","20000714",0,0,0,"PST","15:00:00","16:00:00",416,417,"MASS
_PAR_PU0000010000_BAM_BTD_HD_H1",124,"V0"," ","",20,,"HD",,,,,,,"","",
8,"",207,"EDW","20000714","20000714",0,0,0,"PST","16:00:00","17:00:00",416,417,"MASS
_PAR_PU0000010000_BAM_BTD_HD_H1",124,"V0"," ","",33,,"HD",,,,,,,"","",
8,"",207,"EDW","20000714","20000714",0,0,0,"PST","17:00:00","18:00:00",416,417,"MASS
_PAR_PU0000010000_BAM_BTD_HD_H1",124,"V0"," ","",38,,"HD",,,,,,,"","",
8,"",207,"EDW","20000714","20000714",0,0,0,"PST","18:00:00","19:00:00",416,417,"MASS
_PAR_PU0000010000_BAM_BTD_HD_H1",124,"V0"," ","",37,,"HD",,,,,,,"","",
8,"",207,"EDW","20000714","20000714",0,0,0,"PST","19:00:00","20:00:00",416,417,"MASS
_PAR_PU0000010000_BAM_BTD_HD_H1",124,"V0"," ","",28,,"HD",,,,,,,"","",
8,"",207,"EDW","20000714","20000714",0,0,0,"PST","20:00:00","21:00:00",416,417,"MASS
_PAR_PU0000010000_BAM_BTD_HD_H1",124,"V0"," ","",18,,"HD",,,,,,,"","",
```

Figure 4-3. Example of ARB format showing an excerpt of BAM PM_{10} data collected at EDW. This file shows the Header, notes and a brief sample of data.

5. DATA QUALITY STATEMENTS

5.1 INTRODUCTION AND OBJECTIVES

Data quality summary reports (DQSRs) were prepared for each parameter that STI delivered to the ARB database. The purpose of a DQSR is to provide data users with information on the specifications of data quality. The DQSRs are provided separately (Hyslop et al., 2003). The DQSRs provide the following information

- Operating sites and times for a parameter's measurement
- Data quality objectives (discussed in Section 5.2)
- Data recovery and completeness (discussed in Section 5.3)
- The lower quantifiable limit (LQL) (discussed in Section 5.4)
- Accuracy (discussed in Section 5.5)
- Precision (discussed in Section 5.6)

Table 5-1 lists the measurements, data quality values computed from the data, and the method/data used in the computation.

5.2 DATA QUALITY OBJECTIVES

All data have error, and data analysts need to understand the uncertainties in the data before embarking upon their analyses. Data quality objectives (DQOs) were set for most parameters before the start of CRPAQS. DQOs address goals for completeness, LQL, accuracy, and precision. For some parameters, DQOs were compared to both NARSTO field studies and ARB DQOs.

5.3 DATA COMPLETENESS

Data completeness quantifies the percentage of valid 5-minute, 10-minute, and 60-minute data points. Validity is defined for this calculation as any data point that has a quality control flag of zero. Data completeness is the ratio of the number of valid data points to the number of captured data points.

For each site and parameter, the following quantities were summarized

- The number of sampling periods (5-minute, 10-minute, or 60-minute) in the site-parameter operating date range (the expected number of samples).

- The total number of samples delivered.

- The numbers of valid, invalid, suspect, and missing samples.

Table 5-1. List of measurements, data quality values computed from the data, and the method/data used in the computation.

Measurement	Interval[a]	DQO	Completeness	LQL	Precision computed by calibrations	Precision computed by constant period	Accuracy of flow rates computed	Accuracy computed by span
BC(1,7)[b]	5, 60	✓	✓	✓		✓	✓	
BAM PM_{10}, $PM_{2.5}$	60	✓	✓	✓		✓	✓	
WS, WD, T, RH	5, 60	✓	✓	✓				
Nitric Acid	5, 60	✓	✓	✓	✓			✓
b_{sp}, RH	5, 60	✓	✓	✓	✓		✓	
Nitrate	10, 60	✓	✓	✓	✓	✓		
NO, NO_y	5, 60	✓	✓	✓	✓			✓
Ozone	5, 60	✓	✓	✓	✓			✓
OC, OCEC	60	✓	✓	✓		✓		
PAN, NO_2	5, 60	✓	✓	✓				✓
Climet OPC particle counts	5, 60	✓	✓	✓		✓		
PMS OPC particle counts	5, 60	✓	✓	✓		✓		
SMPS particle counts	5, 60	✓	✓	✓		✓		
SO_2	5, 60	✓	✓	✓	✓			✓
Sulfate	10, 60	✓	✓	✓	✓	✓		

[a] Interval for which data are reported
[b] BC(1,7) = black carbon, 1- and 7-wavelengths

5.4 LOWER QUANTIFIABLE LIMIT

The LQL is the lowest concentration in ambient air that can be measured when processing actual samples. Sources of variability that influence the monitored signal at low concentrations include instrument noise and atmospheric variability. As a measure of this variability, two times the standard deviation of selected 5-minute, 10-minute, and 60-minute data was used to estimate the LQL. The selected data were collected during periods when concentrations were close to the background level and were relatively stable (as measured by a rolling standard deviation). The number of data values and selection criteria varied by pollutant and are discussed in each DQSR. This is a conservative estimate of the LQL because it includes the concentration variability of the ambient air.

The LQL is calculated as shown in Equation 5-1 (for ozone as an example).

$$\mathrm{LQL} \approx 2\sigma = 2\sqrt{\frac{\sum (O_3 - \overline{O}_3)^2}{N-1}} \qquad \textbf{(5-1)}$$

where

$\overline{O}_3$ = mean O_3 concentration
N = number of measurements
σ = standard deviation

The LQL was calculated for all reported time bases (5-, 10-, and 60-minutes) for each parameter.

5.5 ACCURACY

Accuracy is how close the measurements are to the real number. Calibration data are needed to assess accuracy. Accuracy can be evaluated using different approaches. Two techniques are used to evaluate the deviation of measurements from a standard reference. The first method quantifies the variability in the routine accuracy of the instrument by evaluating the span checks, which were performed nightly during CRPAQS. The second method quantifies the accuracy of the instrument response throughout its range of measurement by evaluating the accuracy of multi-point calibration checks. For the Aethalometer, nephelometer, and BAM, only the accuracy of the flow rates could be quantified.

For the gaseous instruments, span checks were performed using the on-site calibrator every night. These nightly checks can be used to evaluate the accuracy of the instrument throughout the study. Accuracy can be expressed in terms of the 95% confidence interval; that is, the reported concentration has a 95% probability of being within ±p of the mean. For surface ozone measurements, for example, the 95% confidence intervals were calculated from the differences between monitor response and known concentrations provided by the automatic span checks performed every night during routine operation. The nightly span checks were performed at 80 ppb. The 95% confidence interval approximates the accuracy of the data as shown in Equation 5-2.

$$\text{Accuracy} \approx 95\% \text{ confidence interval} = \frac{1.96\left(\frac{\sigma_{span}}{\sqrt{N}}\right)}{\left[\overline{O_3}\right]_{cal}} \times 100\% \quad \textbf{(5-2)}$$

where

$$\sigma_{span} = \sqrt{\frac{\sum (x - \bar{x})^2}{N-1}}$$
$$x = [O_3]_{cal} - [O_3]_{measure}$$
$$\bar{x} = \frac{\sum ([O_3]_{cal} - [O_3]_{measure})}{N}$$
$[O_3]_{cal}$ = Ozone concentration output by the calibrator
$[O_3]_{measure}$ = Ozone concentration measured by the analyzer.

The average span concentration measured, $[O_3]_{measure}$, was calculated by averaging the stable span measurements each night. Typically, there were 6 minutes of stable span measurements each night. A small number of span checks were eliminated because the instrument or the calibrator malfunctioned.

5.6 PRECISION

Precision is confidence in the measurement which can be estimated by repeated measurements or duplicate analyses (reproducibility). The consistency of the nightly span concentrations provided a measure of precision in the analyzer measurements. The precision was evaluated by comparing the nightly span concentration to the average span concentration for the entire study. The coefficient of variation (CV) of the span measurements estimates the precision of the data as shown in Equation 5-3 (for ozone as an example).

$$\Pr ecision \approx CV = \frac{\sigma_{measure}}{\left[\overline{O_3}\right]_{measure}} \times 100\% \quad \textbf{(5-3)}$$

where

$$\sigma_{measure} = \sqrt{\frac{\sum \left([O_3]_{measure} - \left[\overline{O_3}\right]_{measure}\right)^2}{N-1}}$$

All the O_3 concentrations in Equation 5-3 refer to the concentrations measured during the span checks.

Precision can be estimated for measurements without span data by evaluating the variance of the pollutant concentrations during periods of low variability when atmospheric influence on variability is assumed to be minimal. Target periods had concentrations well above the LQL. Precision was evaluated as shown in Equation 5-3.

6. REFERENCES

Hansen A.D.A. (2002) Magee Scientific. Personal communication.

Hyslop N.P., Brown S.G., Gorin C.A., and Hafner H.R. (2003) California Regional $PM_{10}/PM_{2.5}$ Air Quality Study data quality summary reports. Final report prepared for San Joaquin Valleywide Air Pollution Study Agency c/o California Air Resources Board, Sacramento, CA by Sonoma Technology, Inc., Petaluma, CA, STI-999242-2310-FR, February.

Kreisberg N. (2002) Personal Communication.

LaRosa L.E., Buckley T.J., and Wallace L.A. (2002) Real-time indoor and outdoor measurements of black carbon in an occupied house: an examination of sources. *J. Air & Waste Manage. Assoc.* **52**, pp. 41-49.

Magee Scientific Company (2002) The Aethalometer™ - real-time aerosol analysis for measuring carbonaceous aerosol species. Available on the Internet at <http://www.mageesci.com>, last accessed January 16, 2003.

Magliano K., Roth P., Blanchard C., Reynolds S., Ziman S., and DeMandel R. (1999) California Regional $PM_{10}/PM_{2.5}$ Air Quality Study: Objectives and associated data analysis and modeling approaches. Draft report prepared by California Air Resources Board, Envair, and Bay Area Air Quality Management District, February.

McDade C.E. (2002) California Regional $PM_{10}/PM_{2.5}$ Air Quality Study (CRPAQS) introduction to site documentation reports. Draft report prepared for California Regional $PM_{10}/PM_{2.5}$ Air Quality Study Technical Committee, c/o California Air Resources Board, Sacramento, CA, by ENSR International, Camarillo, CA, February.

Richards L.W., Lehrman D.E., Weiss R.E., Bush D., Watson J.G., McDade C.M., and Magliano K. (2001a) Light scattering measurements during the California Regional $PM_{10}/PM_{2.5}$ Air Quality Study. In *Regional Haze and Global Radiation Balance - Aerosol Measurements and Models: Closure, Reconciliation, and Evaluation*, S.F. Archer, J.M. Prospero, and J. Core, eds., Air & Waste Management Association, Pittsburgh, PA (STI-2098).

Richards L.W., Weiss R.E., and Waggoner A.P. (2001b) Radiance Research Model 903 integrating nephelometer. In *Regional Haze and Global Balance - Aerosol Measurements and Models: Closure, Reconciliation and Evaluation,* S.F. Archer, J.M. Prospero, and J. Core, eds., Air & Waste Management Association, Pittsburgh, PA (STI-2099).

U.S. Environmental Protection Agency (1980) Validation of air monitoring data. Report prepared by the U.S. Environmental Protection Agency, Research Triangle Park, NC, EPA-600/4-80-030.

Watson J.G., DuBois D.W., DeMandel R., Kaduwela A., Magliano K., McDade C., Mueller P.K., Ranzieri A., Roth P.M., and Tanrikulu S. (1998) Aerometric monitoring program plan for the California Regional $PM_{2.5}/PM_{10}$ Air Quality Study. Draft report prepared for the California Regional $PM_{10}/PM_{2.5}$ Air Quality Study Technical Committee, California Air Resources Board, Sacramento, CA, by Desert Research Institute, Reno, NV, DRI Document No. 9801.1D5, December.

Wittig A.E., Hyslop N.P., and Roberts P.T. (1999) Health and safety plan for the California Regional $PM_{10}/PM_{2.5}$ Air Quality Study. Prepared for the San Joaquin Valleywide Air Pollution Study Agency c/o California Air Resources Board, Sacramento, CA, Sonoma Technology, Inc., Petaluma, CA, STI-999215-1894-HASP, September.

Wittig A.E., Hyslop N.P., and Roberts P. (2000) Quality integrated workplan for the continuous and filter air quality measurements in the California Regional $PM_{10}/PM_{2.5}$ Air Quality Study. Prepared for the San Joaquin Valleywide Air Pollution Study Agency c/o California Air Resources Board, Sacramento, CA, by Sonoma Technology, Inc., STI-999214-1921-QIWP, April.

Wittig A.E., Blumenthal D.L., Roberts P.T., and Hyslop N.P. (2003) California Regional $PM_{10}/PM_{2.5}$ Air Quality Study anchor site measurements and operations. Final report prepared for the San Joaquin Valleywide Air Pollution Study Agency c/o California Air Resources Board, Sacramento, CA, Sonoma Technology, Inc., Petaluma, CA, STI-999231-2332-FR, May.